



# Finite dimensional state representation of physiologically structured populations

Odo Diekmann<sup>1</sup> · Mats Gyllenberg<sup>2</sup> · Johan A. J. Metz<sup>3,4</sup>

Received: 21 August 2018 / Revised: 15 October 2019 / Published online: 21 December 2019  
© The Author(s) 2019

## Abstract

In a physiologically structured population model (PSPM) individuals are characterised by continuous variables, like age and size, collectively called their *i*-state. The world in which these individuals live is characterised by another set of variables, collectively called the environmental condition. The model consists of submodels for (i) the dynamics of the *i*-state, e.g. growth and maturation, (ii) survival, (iii) reproduction, with the relevant rates described as a function of (*i*-state, environmental condition), (iv) functions of (*i*-state, environmental condition), like biomass or feeding rate, that integrated over the *i*-state distribution together produce the output of the population model. When the environmental condition is treated as a given function of time (input), the population model becomes linear in the state. Density dependence and interaction with other populations is captured by feedback via a shared environment, i.e., by letting the environmental condition be influenced by the populations' outputs. This yields a systematic methodology for formulating community models by coupling nonlinear input–output relations defined by state-linear population models. For some combinations of submodels an (infinite dimensional) PSPM can without loss of relevant information be replaced by a finite dimensional ODE. We then call the model ODE-reducible. The present paper provides (a) a test for checking whether a PSPM is ODE-reducible, and (b) a catalogue of *all* possible ODE-reducible models given certain restrictions, to wit: (i) the *i*-state dynamics is deterministic, (ii) the *i*-state space is one-dimensional, (iii) the birth rate can be written as a finite sum of environment-dependent distributions over the birth states weighted by environment independent 'population outputs'. So under these restrictions our conditions for ODE-reducibility are not only sufficient but in fact necessary. Restriction (iii) has the desirable effect that it guarantees that the population trajectories are after a while fully determined by the solution of the ODE so that the latter gives a complete picture of the dynamics of the population and not just of its outputs.

**Keywords** ODE-reducibility · Linear chain trick · Evolutionary system · Input–output system

## 1 Introduction

From the very beginning of community modelling, ordinary differential equations (ODEs) have been its main tool. This notwithstanding the fact that much earlier Euler (1760) and other mathematicians working on population dynamics had already considered age structured models, see (Bacaër 2008, 2011; Gyllenberg 2007) for more information on the history of population dynamics. This probably had two causes, the architects of the initial flurry in the nineteen-twenties and thirties (cf. Scudo and Ziegler 1998) and their successors like MacArthur and May (cf. Kingsland 1995) got their inspiration from the successes of physics, which is dominated by differential equations, and ODEs are rather easier to write down and analyse than e.g. integral equations. However, the assumptions needed to arrive at ODEs generally match biological reality less closely, and give these models more of a toy character: good to get new ideas, but difficult to match in some detail to concrete ecological systems. That for the latter age may well matter also mathematicians know from their immediate experience: few women give birth before the age of ten and while most humans nowadays reach their seventy's anniversary still few live beyond a century. For this reason, many mathematical modellers turned to age as a structuring variable, even in the non-linear realm. However, for ectotherms, that is, all organisms other than mammals and birds, size usually matters far more than age (cf. de Roos and Persson 2013). We have spent considerable effort in the past to develop tools for studying general physiologically structured models in the hope to gradually overcome the surviving endotherm-bias of the modelling community. Yet ODE models remain paramount as didactical tools and for the initial exploration of so far unexplored mechanisms, notwithstanding the disadvantage that in these models individual level mechanisms generally can only be fudged instead of faithfully represented. Given this situation, it becomes of importance to explore in what manner ODE models fit among the physiologically structured ones. Of course, there is the boringly simple embedding of the unrealistic case where individuals indeed have only a single, or at most a few possible states.

**Example 1.1** Consider a size-structured population with individual size (biomass) denoted as  $x$ , starting from a birth size  $x_b$ , individual growth rate  $g(x, E)$ ,  $E$  a resource density, per capita birth rate  $\beta(x, E)$ , and per capita death rate  $\mu(x, E)$ . For such populations, if

$$\frac{x_b \beta(x, E) + g(x, E)}{x} - \mu(x, E) = v(E), \quad (1.1)$$

the population biomass  $B$  per unit of spatial area or volume (below to be abbreviated as just volume) satisfies

$$\frac{dB}{dt} = v(E)B, \quad (1.2)$$

To see this, observe that the left hand side of (1.1) corresponds to the contribution to the change in population biomass by an individual of size  $x$  expressed as fraction of its own biomass. So if we integrate this term over the biomass density over size (and volume), say  $xn(t, x)$ ,  $n$  the numeric (per unit of volume) size density, we get the total change in biomass (per unit of volume), (cf. de Roos et al. 2013).

If moreover

- (i) the per capita contribution to the “consumptive pressure on a resource unit” is a product of an individual’s size and a size-independent functional response based component, say,  $f(E)/E$ ,
- (ii) all other populations in the community are similarly affected only by our focal population’s biomass, and
- (iii) we ourselves are also only interested in this quantity,

then we can for all practical purposes represent our focal population by no more than its biomass.

We call (1.2) an ODE-reduction of the size-structured population model.

The question then naturally arises whether or not this example of ODE-reducibility is essentially the only one, that is, up to coordinate choices, such as in the case of isomorphs not biomass but its scaled cubic root, length. The following example shows this not to be the case.

**Example 1.2** *Daphnia models.* Now let in the wake of (Kooijman and Metz 1984) and (de Roos et al. 1990) size be represented by length, starting from a size  $x_b$  at birth, the growth rate be given by  $g(x, E) = \delta f(E) - \varepsilon x$ , the per capita birth rate by  $\beta(x, E) = \alpha f(E)x^2$ , the per capita death rate by  $\mu$ , and the per capita resource consumption by  $f(E)x^2$ . (This means that individual biomass,  $w$ , scaled to be equal to  $x^3$ , grows as  $3\delta f(E)w^{2/3} - 3\varepsilon w$ , that is, mass intake is taken to be proportional to surface area and metabolism to biomass.) Let  $n(t, x)$  again denote the numeric size-density. Now define

$$N_i(t) = \int_{x_b}^{x_{\max}} x^i n(t, x) dx, \quad (1.3)$$

that is,  $N_0$  is the total population size,  $N_1$  the total population length,  $N_2$  the total population surface area,  $N_3$  the total population biomass, all per unit of volume. Then

$$\begin{aligned} \frac{dN_0}{dt} &= \alpha f(E)N_2 - \mu N_0, \\ \frac{dN_1}{dt} &= x_b \alpha f(E)N_2 + \delta f(E)N_0 - (\mu + \varepsilon)N_1, \\ \frac{dN_2}{dt} &= x_b^2 \alpha f(E)N_2 + 2\delta f(E)N_1 - (\mu + 2\varepsilon)N_2, \\ \frac{dN_3}{dt} &= x_b^3 \alpha f(E)N_2 + 3\delta f(E)N_1 - (\mu + 3\varepsilon)N_3, \end{aligned} \quad (1.4)$$

etc.. If the only other component of the community is an unstructured resource, and we need no further output from the population than its total biomass per unit of volume, we can combine (1.4) with

$$\frac{dE}{dt} = h(E) - f(E)N_2 \quad (1.5)$$

into a sufficient representation of our community model.

The differential equation for  $N_0$  is obvious, and so is the first term in the differential equation for  $N_1$ . To understand the second term observe that  $g$  consists of two terms, the first of which is size independent. This term makes all individuals of the population increase their length at a rate  $\delta f(E)$ . We get the corresponding increase in the total population length by multiplying this  $\delta f(E)$  with the total population density. The  $\varepsilon$  in the last term also derives from the length growth except that the corresponding term in  $g$  contains a factor  $x$ . When we account for this  $x$  when calculating the integral over  $n$  we get  $N_1$ . The differential equations for the other  $N_i$  follow in a similar manner.

ODE-reducibility of age-structured models and, slightly more generally, of distributed delay systems, has been investigated since the mid 1960s (Vogel 1965; Fargue 1973, 1974; Gurtin and MacCamy 1974, 1979a, b; McDonald 1978, 1989).

It has already been known for a long while that there also exist more realistic cases, where for instance a size-structured model allows a faithful representation in ODE terms (Murphy 1983; Cushing 1989; Metz and Diekmann 1991).

The next question is then whether we can characterise the set of *all* possible cases. For the practically important subset of cases where the population birth rate figures on the list of population outputs and with a single state variable on the level of the individuals, the last author solved this problem on a heuristic level already in 1989 during a holiday week in summer spent at the Department of Applied Physics of the University of Strathclyde. An allusion to this was given in a “note added in print” to the paper (Metz and Diekmann 1991). However, it took till now before we together had plugged all the minor holes in the proof. Below you find the result.

## 2 Preview of Sects. 3 to 6

In this section we give a preview of the main content of the paper, first for theoretical biologists and probabilists and then for all kinds of mathematicians. The much shorter paper (Diekmann et al. 2019) provides additional examples and may serve as a more friendly user guide to ODE-reducibility of structured population models.

### 2.1 Mainly for theoretical biologists and probabilists: the context of discovery

#### 2.1.1 Biological context

The term “physiologically structured population models” (PSPM) refers to large system size limits of individual-based models where (i) individuals are differentiated by physiological states, e.g.  $x = (\text{size, age})$ , referred to as i(ndividual)-states, (ii) the

world in which these individuals live is characterized by a set of variables collectively called environmental condition, to be denoted as  $E$ . [Hard proofs for the limit conjectures are still lacking except for age-based models (cf. Tran 2008, and the references therein), and more recently also for a class of simple (size,age)-based ones (Metz and Tran 2013).] Sections 2.2 and 3 go into how these deterministic models can be specified by means of equations.

The i-level model ingredients are a set of feasible i-states,  $\Omega \subset \mathbb{R}^m$ ,  $m \in \mathbb{N}$ , and, most commonly,

- (i) a rate of i-state change taken to be deterministic,  $g(x, E)$ ,
- (ii) a death rate,  $\mu(x, E)$ , and
- (iii) a birth rate,  $\beta(x, E)$ .

In the general case  $\beta(x, E)$  has as value a distribution over  $\Omega$ . However, from a mathematical perspective it is preferable to use instead of “distribution” the term “measure” as this is more encompassing, and the birth rate does not total to one and may consist of a mixture of discrete and continuous distributions. (Actually, we should even be a bit more general and talk about a signed measure as a cell that divides generates a measure over the states where the daughters may land plus a compensating negative mass, equal to minus the division rate, at the state of the mother.)

*Notational convention* The value of  $\beta(x, E)$  for the measurable set  $\omega \subset \Omega$  is denoted by  $\beta(x, E, \omega)$ . A similar convention applies to other measure valued functions.

The p(opulation)-state then is a measure  $m$  on  $\Omega$ . However, on many occasions it suffices to think in terms of just densities  $n$  on  $\Omega$ , or  $n \in L^1(\Omega)$  in the mathematicians’ jargon. As a consequence of how the rates are specified, when  $E$  is given as a function of time (below to be looked at as input) the individuals are independent (except for a possible dependence of their birth state on the state of their parents), and hence the dynamics of the p-state is linear.

The more interesting case is when  $E$  is determined by the surrounding community. Community models are sets of population models coupled through a common  $E$ . This leads to c(ommunity)-state spaces that are products of the state spaces of the comprised species, times the state spaces of any non-living resources. The mass action principle tells that generally  $E$  can be calculated by applying a linear map to the c-state, like when a predation pressure equals the sum of the predation pressures exerted by all individuals in the community. This leads us to the final set of ingredients of a population model:

- (iv) functions of  $(x, E)$ , like biomass or per capita feeding rate, that when cumulated over all individuals produce components of the population output.

Side remarks on terminology: In our context, each output component is thus obtained by taking the integral of the p-state over  $\Omega$  after multiplying it with a, possibly  $E$ -dependent, function of the i-state variables. This function specifies the relationship between the i-state and the property that we want to measure, e.g., biomass as a function of length. In mathematical jargon we say that the output components are obtained by applying a linear functional, i.e., a linear map from the p-state space to the real numbers. The corresponding function will be referred to as weight function, and for a p-state  $m$  and weight function  $\psi$ , the corresponding map will be written as  $\langle m, \psi \rangle$ .

The population dynamical behaviour of individuals almost never can be captured in terms of a finite number of  $i$ -states. Yet, ecological discourse is dominated by ODE models. This leads to the philosophical question which of these models can be justified from the more realistic physiologically structured populations perspective. At the more pragmatic side there is moreover the problem that in community biology PSPM usually become too difficult to handle for more than two or three species. This leads to the complementary question whether there are relevant choices of model ingredients for which a PSPM can be represented by a low dimensional system of ODEs. To answer these questions we looked at populations as state-linear input–output relations, with  $E$  as input, and as output a population's contribution to  $E$  as well as anything that 'a client' may want to keep track of. The key question addressed in this paper is thus: under what conditions on the model ingredients is it possible to obtain the same input–output relation when the PSPM is replaced by a finite dimensional ODE? (This representation may have an interpretation in its own right, but this is not required.) If such a representation is possible, we say that the population model is reduced to the ODE or that the input–output relation is realised by it.

### 2.1.2 The mathematical question

Our starting point thus are models that can be represented as in the following diagram. In Fig. 1  $Y$  is the  $p$ -state space and  $\mathbb{R}^r$  the output space.  $E$  is the time course of the environment and  $U_E^c(t, s)$  the (positive) linear state transition map with  $s, t$  the initial and final time. (The upper index  $c$  here refers to the mathematical construction of the  $p$ -state, explained in Sect. 3, through the cumulation of subsequent generations.) Finally  $O(E(t))$  is the linear output map. The mathematical question then is under which conditions on the model ingredients it is possible to extend the diagram in Fig. 1 (for all  $E, t, s$ ) to the following diagram.

Here  $P$  is a linear map,  $\Phi_E(t, s)$  a linear state transition map (which should be differentiable with respect to  $t$ ) and  $Q(E(t))$  a linear output map. The dynamics of the output cannot be generated by an ODE when the space spanned by the output vectors at a given time is not finite dimensional. Hence ODE reducibility implies that there exists an  $r$  such that the outputs at a given time can be represented by  $\mathbb{R}^r$ . (Below we drop the time arguments to diminish clutter, except in statements that make sense only for each value of the argument separately, or when we need to refer to those arguments.) Moreover, the biological interpretation dictates that

$$O(E)m = \langle m, \Gamma(E) \rangle := \int_{\Omega} \Gamma(E)(x)m(dx),$$

where  $m$  is the  $p$ -state and the components of the vector  $\Gamma(E)$  are functions  $\gamma_i(E) : \Omega \rightarrow \mathbb{R}$ .

$$Y \xrightarrow{U_E^c(t, s)} Y \xrightarrow{O(E(t))} \mathbb{R}^r$$

Fig. 1 Structure of models with output

Thanks to the linearity of  $U_E^c(t, s)$  and  $O(E(t))$  we can without loss of generality assume  $P$ ,  $\Phi_E(t, s)$  and  $Q(E(t))$  to be linear. Moreover, ODE reducibility requires that  $P$  can be written as  $Pm = \langle m, \Psi \rangle$  with  $\Psi = (\psi_1, \dots, \psi_k)^T$ ,  $\psi_i : \Omega \rightarrow \mathbb{R}$ , where the  $\psi_i$  should be sufficiently smooth to allow

$$dN/dt = K(E)N, \text{ with } N := Pm \text{ and } K(E(t)) := d\Phi_E(t, s)/dt|_{s=t}. \quad (2.1)$$

(The last expression comes from combining  $d\Phi_E(t, s)/dt = K(E(t))\Phi_E(t, s)$  and  $\Phi_E(t, t) = I$ .) Finally, we should have  $O(E) = Q(E)P$ , and therefore  $\Gamma(E) = Q(E)\Psi$ , implying that the output weight functions should be similarly smooth. (The precise degree of smoothness needed depends on the other model ingredients in a manner that is revealed by the TEST described in Sect. 2.1.4.)

### 2.1.3 A tool

The main tool in the following considerations is the so-called backward operator, to be called  $\bar{A}(E)$ , as encountered in the backward equation of Markov process theory and thereby in population genetics, used there to extract various kinds of information from the process without solving for its transition probabilities. What matters here is that the backward operator summarises the behaviour of the “clan averages”  $\bar{\psi}(t, s)$  of weight functions  $\psi$  (such as occur in  $Pm = \langle m, \Psi \rangle$ ), defined by

$$\bar{\psi}(t, s)(x) := \int_{\Omega} \psi(\xi) m(t, s; \delta_x, d\xi), \quad (2.2)$$

where  $m(t, s; \delta_x)$  is the p-state resulting at  $t$  from a p-state corresponding to a unit mass  $\delta_x$  at  $x$  at time  $s$  (hence the term clan average), in the sense that

$$\bar{A}(E(s))\bar{\psi}(t, s) = -\frac{\partial \bar{\psi}(t, s)}{\partial s}. \quad (2.3)$$

For further use we moreover note that

$$\bar{\psi}(s, s) = \psi, \quad (2.4)$$

independent of  $s$ , which on differentiating for  $s$  gives

$$\text{at } t = s : -\frac{\partial \bar{\psi}(t, s)}{\partial s} = \frac{\partial \bar{\psi}(t, s)}{\partial t}. \quad (2.5)$$

(When  $E$  does not depend on  $s$ , (2.5) also holds good for  $t \neq s$ , leading to the perhaps more familiar form of the backward equation:  $d\bar{\psi}/dt = \bar{A}\bar{\psi}$ .)

One reason to start from (2.2) is that it leads to simple interpretation-based heuristics for calculating backward operators, which we shall discuss a little later on. (More abstract and rigorous functional analytical counterparts to the intuitive interpretation-based line of argument developed here can be found in Sects. 2.2 and 4.) The reason for

coming forth with the backward operator is that it provides the counterpart in the space of weight functions  $\psi$  of the time differentiation in (2.1). To see that, the perspective sketched so far still needs a slight extension. A look at (2.2) shows that we can also interpret the  $\bar{\psi}(t, s)$  as weight functions in their own right that can be paired with  $\delta_x$  as  $\langle \delta_x, \bar{\psi}(t, s) \rangle (= \bar{\psi}(t, s)(x))$ . Moreover, thanks to the linearity of (2.2), (2.3) and the p-state transition maps  $U_E^c(t, s)$  and the consequent linearity of  $m(t, s; \cdot)$ , we can extend the action of these candidate functionals to more general measures  $m_s$  at time  $s$ , and in this way calculate  $\int_{\Omega} \psi(\xi) m(t, s; m_s, d\xi)$  as  $\langle m_s, \bar{\psi}(t, s) \rangle$ . This sleight of hand transforms the question about the change of  $Pm$  over time to a question about the dynamics of the  $\bar{\psi}(t, s)$ , for which we can find the time derivative by applying  $\bar{A}(E)$  to  $\bar{\psi}$ . The final step then is to make the connection with (2.1) by setting  $s = t$  and using (2.4), to arrive at

$$\bar{A}(E)\Psi = K(E)\Psi. \quad (2.6)$$

To apply these ideas we need to find expressions for the backward operators. To this end we use that we assumed  $E$  to be given, making individuals independent. Then on the level of population averages it makes no difference whether we start with a scaled large number of individuals all with i-state  $x$  or with a single individual in that state. Since it is easier to think in terms of individuals, we shall do the latter. We can then look at the effect on the clan averages of the founder individual engaging in each of the component processes (i) to (iii) over a short time interval from  $s$  to  $s + h$ . For such short intervals the effects of the interaction of the components in determining their combined effect on the clan average is only second order in  $h$  and can be neglected together with the other higher order terms. Hence the backward operator can be written as a sum of operators corresponding to the model ingredients as introduced at the beginning of Sect. 2.1.1. As this is useful for the remainder of the paper we here combine ingredients (i) and (ii):

$$\begin{aligned} \bar{A}(E) &= \bar{A}_0(E) + \bar{B}(E), \\ \text{(i) and (ii): } (\bar{A}_0(E)\psi)(x) &= \frac{d\psi}{dx}(x)g(x, E) - \mu(x, E)\psi(x), \\ \text{(iii): } (\bar{B}(E)\psi)(x) &= \int_{\Omega} \psi(\xi)\beta(x, E, d\xi). \end{aligned} \quad (2.7)$$

The expressions in (2.7) are found as follows:

- (i) If we neglect births and deaths, for small  $h$  an individual situated at  $x$  at time  $s$  produces an individual situated at  $x + g(x, E(s))h$  at time  $s + h$ . Therefore  $\bar{\psi}(t, s + h)(x + g(x, E(s))h) = \bar{\psi}(t, s)(x)$ , which after on both sides subtracting  $\bar{\psi}(t, s + h)(x)$  and dividing by  $h$ , gives

$$-\frac{\partial \bar{\psi}(t, s)(x)}{\partial s} \Big|_{\text{movement}} = \frac{\partial \bar{\psi}(t, s)(x)}{\partial x} g(x, E(s)). \quad (2.8)$$



- (ii) Next we account for the possibility that individuals die on the way, which happens with probability  $\mu(x, E(s))h$ , so that, if we neglect movement and births,  $\bar{\psi}(t, s)(x) = (1 - \mu(x, E(s))ds)\bar{\psi}(t, s + h)(x)$  (over a longer time survival is less), giving

$$-\left. \frac{\partial \bar{\psi}(t, s)(x)}{\partial s} \right|_{\text{death}} = -\mu(x, E(s))\bar{\psi}(t, s)(x). \quad (2.9)$$

- (iii) Finally a parent at  $x$  through its kids at  $\xi$  will have added a contribution  $h\beta(x, E(s), d\xi)\bar{\psi}(t, s)(\xi)$  to  $\bar{\psi}(t, s)(x)$ , which is missing in  $\bar{\psi}(t, s + h)(x)$ . Summing over all these contributions gives

$$-\left. \frac{\partial \bar{\psi}(t, s)(x)}{\partial s} \right|_{\text{births}} = \int_{\Omega} \bar{\psi}(t, s)(\xi)\beta(x, E(s), d\xi). \quad (2.10)$$

### 2.1.4 Testing combinations of model ingredients for ODE reducibility

The classical search heuristic for finding a state space representation for a given set of dynamical variables in continuous time is to see whether their derivatives can be expressed in terms of the variables themselves, and, if not, to join the derivatives for which this is not the case as additional prospective state variables to the original set, whereupon the procedure is repeated till one succeeds or runs out of patience (cf. Diekmann et al. 2018, p. 1443; Fargue 1973).

If one is after a representation as a state linear system with a linear output map, “can be expressed as” translates to “is linearly dependent on”, a property that can be checked algorithmically, so that with infinite patience we end up with a firm conclusion. In our case the only difference is that we choose not to look at the output variables themselves, but at the weight functions by which they are produced from the population state, and therefore replace the time derivative of the output variables by the backward operator applied to these functions.

*Notation* For the remainder of this subsection we shall again use  $E$  to denote the environmental conditions, as opposed to the course of the environment as we did in the previous two subsections.

#### TEST

Choose a basis  $V_0 = \{\psi_1, \dots, \psi_{k_0}\}$  for the  $\gamma_i$  making up the population output map.

Provided that the expression  $\bar{A}(E)V_i$  makes sense, extend  $V_i$  to a basis  $V_{i+1}$  for  $V_i \cup \bigcup_{\text{all possible } E} \bar{A}(E)V_i$ .

The population model is ODE reducible if and only if the expressions  $\bar{A}(E)V_i$  keep making sense, and the  $V_i$  stop increasing after a certain  $i = h$ .

For the  $\bar{A}(E)$  from (2.7),  $\bar{A}(E)V_i$  makes sense iff all elements of  $V_i$  are differentiable. (For a mathematically more precise rendering see Sect. 4.)

If the biological ingredients of a model satisfy the above test, it is possible to go directly from the ingredients to the ODE.

**Example 2.1** *Daphnia models, continued.* In Example 1.2 the output weight functions were  $x^3$  and  $x^2$ . Applying the backward operator

$$\bar{A}(E)\psi(x) = (\delta f(E) - \varepsilon x) \frac{d\psi}{dx}(x) - \mu\psi(x) + \psi(x_b)\alpha f(E)x^2$$

to these functions gives

$$\begin{aligned} \bar{A}(E) \begin{pmatrix} x^3 \\ x^2 \end{pmatrix} &= \begin{pmatrix} (\delta f(E) - \varepsilon x)3x^2 - \mu x^3 + x_b^3 \alpha f(E)x^2 \\ (\delta f(E) - \varepsilon x)2x - \mu x^2 + x_b^2 \alpha f(E)x^2 \end{pmatrix} \\ &= \begin{pmatrix} x_b^3 \alpha f(E)x^2 + 3\delta f(E)x^2 - (\mu + 3\varepsilon)x^3 \\ x_b^2 \alpha f(E)x^2 + 2\delta f(E)x - (\mu + 2\varepsilon)x^2 \end{pmatrix}. \end{aligned} \quad (2.11)$$

This introduces  $x$  as an additional weight function, linearly independent of  $x^3$  and  $x^2$ . Applying the backward operator to  $x$  gives

$$\bar{A}(E)x = \delta f(E) - \varepsilon x - \mu x + x_b \alpha f(E)x^2 = x_b \alpha f(E)x^2 + \delta f(E) - (\mu + \varepsilon)x, \quad (2.12)$$

introducing the constant 1 as further weight function. Applying  $\bar{A}(E)$  to 1 gives

$$\bar{A}(E)1 = -\mu + x_b \alpha f(E)x^2 = x_b \alpha f(E)x^2 - \mu, \quad (2.13)$$

which introduces no further linearly independent weight functions, and thus ends the process. Hence, a  $P$  can be built from the weight functions 1,  $x$ ,  $x^2$  and  $x^3$  leading to the four dimensional ODE representation already derived in Example 1.2.

Integrating (2.11) to (2.13) left and right over the p-state gives (1.4). This is why we could reorder (2.11) to (2.13) to look like (1.4), with the  $N_i$  replaced by  $x^i$ . Rewritten in matrix form this then gives the  $K(E)$  from (2.1). Note though that in the process we have made some particular choices in order to get the precise form (1.4). In general  $K(E)$  is only unique up to a similarity transform, corresponding to alternative choices of a basis for the weight functions  $\psi_1, \dots, \psi_k$ , with a corresponding change of  $Q(E)$ .

**Remark** At first sight the TEST may not seem very practical as deciding that a certain set of model ingredients is not ODE reducible may require infinitely many operations. However, if the model ingredients come in the form of explicit expressions it can often be inferred from these whether the combination of ingredients will or will not pass the test. And even when those expressions on first sight are less than transparent, not all is lost, as in practice people are generally not so much interested in whether a certain model is ODE reducible as such, but in whether there exists a representation with a relatively low dimensional state space, which after specification of the maximum allowed dimension leads to a task executable by e.g. Maple<sup>TM</sup> or Mathematica<sup>®</sup>.

## 2.1.5 A catalogue of ODE reducible models

Modellers often initially still have quite some freedom in their choice of model ingredients. Hence it is useful to construct catalogues of classes of ODE reducible models,

to help them make their choices with an eye on future model tractability (for a good example, see (Cushing 1989)). For this purpose observe that the TEST is no more than an attempt to construct a vector  $\Psi$  of weight functions such that for all environmental conditions  $E$  (i)  $\Gamma(E) = Q(E)\Psi$  and (ii)

$$\bar{A}(E)\Psi = K(E)\Psi, \quad (2.14)$$

suggesting the following strategy: start from some specific promising model class, and within this class try to solve (2.14) for  $\Psi$  and  $K$ . In Sects. 5 and 6 we concentrate on one particular class of models that is both hallowed by tradition, and has the good property of allowing the p-state trajectory to be reconstructed by relatively easy means:

$$\beta(x, E) = \sum_{i=1}^p \beta_i(E)\alpha_i(x) \quad (2.15)$$

with the  $\alpha_i$  treated as output weight functions au par with the other  $\gamma_i$ . (For the numerics it helps when the  $\beta_i(E)$  consist of just a few point masses at positions that depend smoothly on  $E$ .) For the remainder of this subsection we assume that (2.15) holds good.

**Remark** In earlier papers like (Metz and Diekmann 1991) we referred to ODE reducible models satisfying (2.15) as ordinarily ODE reducible or ODE reducible *sensu stricto*, as these models then were the only ones figuring in discussions of ODE reducibility (or linear chain trickability as it then was called).

The assumption that the  $\alpha_i$  are additional output weight functions makes that we can absorb the birth term of the backward operator into the right hand side of (2.14) giving

$$\bar{A}_0(E)\Psi = H(E)\Psi. \quad (2.16)$$

All our general results so far pertain to the case where  $\dim(\Omega) = 1$ . In Sect. 8 we show how the results for this case can be combined to make ODE reducible models with  $\dim(\Omega) > 1$ , but a full catalogue for general  $\Omega$  is still lacking. So for the remainder of this subsection we assume that  $\dim(\Omega) = 1$ .

Below we give a gross sketch of the reasoning. As first step we choose a fixed constant  $E = E_r$ , write  $v_r(x) = g(x, E_r)$ ,  $\mu_r(x) = \mu(x, E_r)$ ,  $H_r = H(E_r)$ , where the label r in the subscripts stands for “reference”, and solve the corresponding version of (2.16),

$$\frac{d\Psi}{dx}(x)v_r(x) - \mu_r(x)\Psi(x) = H_r\Psi(x), \quad (2.17)$$

for  $\Psi$ . The result is that  $\Psi(x)$  is the product of

$$\mathcal{G}_r(x) := \exp\left(\int^x \frac{\mu_r(\xi)}{v_r(\xi)} d\xi\right)$$

(note that  $\mathcal{G}_r(x)$  can be interpreted as the inverse of a survival probability up to size  $x$ ) and a matrix exponential in the transformed  $i$ -state variable

$$\zeta_r(x) := \int^x \frac{d\xi}{v_r(\xi)}.$$

This tells that the weight functions  $\psi$  should be linear combinations of polynomials times exponentials in  $\zeta_r$ , all multiplied with the same  $\mathcal{G}_r$ . The fact that these weight functions should not depend on  $E$  then gives a set of conditions that should be satisfied by the  $\Psi$ ,  $g$  and  $\mu$  together. There are three possibilities to satisfy these conditions:

(i) and (ii) Without any restrictions on  $H_r$ , the death rate  $\mu$  should be decomposable as  $g(x, E)\gamma_0(x) + \mu_0(E)$ , that is, a death rate component that at each value of  $x$  depends on how fast an individual grows through this value plus a death rate component that does not depend on  $x$ , and

- (i) the representation should be one-dimensional, in which case  $\Psi(x) = d\mathcal{G}(x)$  with  $\mathcal{G}(x) := \exp\left(\int^x \gamma_0(\xi)d\xi\right)$ ,  $d$  a scalar of choice, and  $H(E) = \mu_0(E)$ , or, in the higher dimensional case,
- (ii)  $g$  should be decomposable as  $g(x, E) = v_0(x)v_1(E)$ , so that we may interpret  $\zeta(x) := \int^x (v_0(\xi))^{-1}d\xi$  as physiological age.  $H$  in this case can be written as  $H(E) = -\mu_0(E)I + v_1(E)L$ , with  $L$  an arbitrary matrix with all eigenvalues having geometric multiplicity one.  $\Psi$  can then be written as

$$\Psi(x) = \mathcal{G}(x)D \left( e^{\lambda_1 \zeta}, \dots, \zeta^{k_1-1} e^{\lambda_1 \zeta}, \dots, e^{\lambda_r \zeta}, \dots, \zeta^{k_r-1} e^{\lambda_r \zeta} \right)^T,$$

$D$  a nonsingular matrix, with a corresponding representation of  $L$  as

$$L = D \begin{pmatrix} \begin{pmatrix} \lambda_1 & & 0 \\ 1 & \ddots & \\ & \ddots & \ddots \\ 0 & & k_1-1 & \lambda_1 \end{pmatrix} & & 0 \\ & \ddots & \\ & & 0 & \begin{pmatrix} \lambda_r & & 0 \\ 1 & \ddots & \\ & \ddots & \ddots \\ 0 & & k_r-1 & \lambda_r \end{pmatrix} \end{pmatrix} D^{-1}.$$

- (iii) For higher dimensional representations also slightly less restricted growth laws are possible, but at the cost of a severe restriction on the eigenvalues of  $H_r$  which should lie in special regular configurations in the complex plane. A lot of hard work is then required to render the corresponding class of representations into a biologically interpretable form. The end result becomes a slight extension of the Daphnia model of Examples 1.2 and 2.1 with  $\mu(x, E) = g(x, E)\gamma_0(x) + \mu_0(E)$ ,

$g(x, E) = v_0(x)(v_1(E) + v_2(E)\zeta)$  (note that if we transform from  $x$  to  $\zeta$  the growth law stays the same but for the disappearance of the factor  $v_0(x)$ ). In this case  $\Psi(x) = \mathcal{G}(x)D\Xi(x)$  with  $\Xi(x) = (1, \dots, \zeta^{k-1})^T$ . (The growth laws of this model family, in addition to von Bertalanffy growth, also encompass two other main growth laws from the literature: logistic,  $a(E)x + b(E)x^2$ , and Gompertz,  $a(E)x - b(E)x \ln(x)$ . Just appropriately transform the  $x$ -axis.)

A further extension comes from a mathematical quirk for which we failed to find a biological interpretation: it is possible to add a quadratic term in  $\zeta$  to the growth law,  $g(x, E) = v_0(x)(v_1(E) + v_2(E)\zeta + v_3(E)\zeta^2)$ , which then should be exactly compensated by a similar additional term in the death rate,  $\mu(x, E) = g(x, E)\gamma_0(x) + \mu_0(E) + (k-1)v_3(E)\zeta$ . Not only are these terms uninterpretable, the required simultaneous fine tuning of the model components makes them irrelevant in any modelling context coming to mind.

For this model family  $H(E) = -\mu_1(E)I + DL(E)D^{-1}$  with

$$L(E) = \begin{pmatrix} 0 & -(k-1)v_3(E) & 0 & \cdots & 0 \\ v_1(E) & v_2(E) & -(k-2)v_3(E) & & 0 \\ 0 & 2v_1(E) & 2v_2(E) & & 0 \\ \vdots & 0 & 3v_1(E) & & 0 \\ & & & \ddots & \\ 0 & \cdots & \cdots & (k-1)v_1(E) & (k-1)v_2(E) \end{pmatrix}$$

with  $v_3 = 0$  in the Daphnia style models.

The matrix  $K$  occurring in the ODE realising the population input–output relation can be recovered by removing the effect of subtracting the birth operator on both sides of (2.14) to get (2.15):  $K(E) = H(E) + \int_{\Omega} \psi_i(\xi)\beta_j(E, d\xi)C$ , with  $C$  defined by the requirement that  $\alpha_j = \sum_h c_{jh}\psi_h$ , where the  $\alpha_j$  are the weight functions telling how the birth rates depend on the  $i$ -states (see Formula (2.15)).

The remainder of the paper is geared to an audience of analysts, and accordingly stresses proofs, instead of interpretation-based heuristics.

## 2.2 For mathematicians: the context of justification

We shall look at a community as a set of coupled populations. The coupling is mediated by the environment, denoted by  $E$ . On the one hand individuals react to the environment, on the other hand they influence their environment. We concentrate on a single population and pretend that  $E$  is a given function of time taking values in a set  $\mathcal{E}$ . So  $E$  can be regarded as an input. The single population model can serve as a building block for the community model once we have also specified a population level output by additively combining the impact that individuals have.

In order to account for the population dynamical behaviour of individuals (giving birth, dying, impinging on the environment), we first introduce the concept of *individual state* ( $i$ -state). Given the course of the environment, the  $i$ -states of the individuals independently move through the  $i$ -state space and their current position is all that matters at any particular time. The use of the word *state* entails the Markov property;

admittedly idealisation is involved and finding the i-state space as well as specifying the relevant environmental variables is often a process of trial and error. The art of modelling comprises deliberate simplification in order to gain significant insights.

We denote the i-state space by  $\Omega$  and assume that it is a subset of  $\mathbb{R}^n$ . In general the *population state* (p-state) is a measure  $m$  on  $\Omega$  with the interpretation that  $m(\omega)$  is the number (per unit of volume or area) of individuals with i-state in the measurable subset  $\omega$  of  $\Omega$ .

*A word of warning* We usually denote a Borel subset of  $\Omega$  by  $\omega$ . We are aware of a different notational convention in probability theory where  $\omega$  denotes a *point* in  $\Omega$ . We hope this will not lead to confusion.

In many models the p-state can adequately be represented by a density  $n \in L^1(\Omega)$ . The abstract ODE

$$\frac{dn}{dt}(t) = A(E(t))n(t) = A_0(E(t))n(t) + B(E(t))n(t) \quad (2.18)$$

captures that the density  $n(t)$  changes in time due to

- (i) transport through  $\Omega$  due to i-state development such as growth of individuals, and degradation due to death of individuals,
- (ii) reproduction.

The effects of (i) are incorporated in the action of  $A_0(E)$  and the effects of (ii) in the action of  $B(E)$ . Since the i-states of offspring are, as a rule, quite different from the i-states of the parent, the operator  $B(E)$  is usually non-local. When  $x \in \Omega$  specifies the size of an individual, growth is deterministic and giving birth does not affect the parent's size, the abstract ODE (2.18) corresponds to the PDE

$$\begin{aligned} \frac{\partial}{\partial t} n(t, x) = & -\frac{\partial}{\partial x} (g(x, E(t))n(t, x)) - \mu(x, E(t))n(t, x) \\ & + \int_{\Omega} \tilde{\beta}(y, E(t), x)n(t, y)dy, \end{aligned} \quad (2.19)$$

with  $g$ ,  $\mu$  and  $\int_{\Omega} \tilde{\beta}(\cdot, \cdot, x)dx$  denoting, respectively, the growth, death, and reproduction rate of an individual with the specified i-state under the specified environmental condition.

The mathematical justification of (2.18) or (2.19) is cumbersome. In particular, it is difficult to give a precise characterisation of the domains of the various unbounded operators. Note that we are not in the setting of (generators of) semigroups of linear operators for which some results can be found in the paper by Atay and Roncoroni (2017). Indeed, we are dealing with evolutionary systems and for this non-autonomous analogue of semigroups a one-to-one correspondence between an evolutionary system and a generating family of differential operators is not part of a well-established theory (but see our earlier work (Clément et al. 1988; Diekmann et al. 1995) for some results in that direction). As shown by Diekmann et al. (1998, 2001), one can actually avoid the differential operators. In the next section we present this approach in the setting where, because of assumptions concerning reproduction, we can work with densities rather

than measures. In the present section we simply ignore the mathematical difficulties and proceed formally.

If the test described in Sects. 4.1 and 2.1.4 yields a positive result in the end, it provides

- an integer  $k$ ,
- a bounded linear map  $P : L^1(\Omega) \rightarrow \mathbb{R}^k$ ,
- a family  $K(E)$  of  $k \times k$  matrices,

such that

$$PA(E) = K(E)P \quad (2.20)$$

and accordingly

$$N(t) := Pn(t), \quad (2.21)$$

where  $n(t)$  is a solution of (2.18), satisfies the ODE

$$\frac{d}{dt}N(t) = K(E(t))N(t). \quad (2.22)$$

The p-output that is needed in the community model, which is our ultimate interest, as well as any other p-output that we are interested in, was the starting point for the test and thus is incorporated in  $N$ , so we can focus our attention on the finite dimensional ODE (2.22) and forget about the infinite dimensional version (2.18) from which it was derived.

A special case occurs when there exist

- a family  $H(E)$  of  $k \times k$  matrices,
- a family  $Q(E)$  of bounded linear maps from  $\mathbb{R}^k$  to  $L^1(\Omega)$

such that

$$PA_0(E) = H(E)P, \quad (2.23)$$

$$B(E) = Q(E)P. \quad (2.24)$$

(Incidentally, we here took  $L^1(\Omega)$  as the range space for  $Q(E)$ , but actually we shall allow  $Q(E)$  to take on values in a linear subspace of the vector space of all Borel measures on  $\Omega$ , see Sect. 4.) This case amounts to taking

$$K(E) = H(E) + M(E). \quad (2.25)$$

with

$$M(E) = PQ(E). \quad (2.26)$$

The key nice features of this special case are:

- (i)  $A_0(E)$  is a strictly local operator and this allows us to make an in-depth study of the characterisation of those  $A_0(E)$  for which  $P$  and  $H(E)$  exist.
- (ii)  $B(E)n = Q(E)N$  and hence we can rewrite (2.18) in the form

$$\frac{dn}{dt} = A_0(E)n + Q(E)N. \quad (2.27)$$

The same local character of  $A_0(E)$  now allows us to solve (2.27) explicitly, thus expressing the part of the population that is born after the time at which we put an initial condition explicitly in terms of  $N(t)$ . See Proposition 7.2 for a concrete example.

If we ignore birth, that is, set  $B(E) = 0$  ( $\tilde{\beta}(\cdot, E, \cdot) = 0$ ), then the abstract ODE (2.18) and its PDE counterpart (2.19) become transport-degradation equations and we only have to consider condition (2.23). In this paper we give necessary and sufficient conditions in terms of  $g$  and  $\mu$  for the existence of  $k$ ,  $P$  and  $H(E)$  such that (2.23) holds and hence  $N(t)$  satisfies (2.22) (with  $K(E) = H(E)$ , or, equivalently  $M(E) = 0$ ). Subsequently we view condition (2.24) as a *restriction* on the submodel for reproduction. If  $\tilde{\beta}$  is such that the restriction (2.24) holds, then the full infinite-dimensional system (2.18) is reducible to the finite dimensional ODE (2.22). In this manner we obtain a catalogue of models that are ODE-reducible and within a restricted class of models with one-dimensional  $i$ -state space the catalogue is even complete.

Let  $w_i$   $i = 1, 2, \dots, k$  be bounded measurable functions defined on  $\Omega$  such that

$$(Pn)_i = \langle n, w_i \rangle = \int_{\Omega} n(x)w_i(x)dx. \quad (2.28)$$

Then condition (2.23) amounts to

$$A_0^*(E)w_i = \sum_{j=1}^k H_{ij}(E)w_j. \quad (2.29)$$

We might call  $A_0^*(E)$  the Kolmogorov backward operator although strictly speaking that operator acts on the continuous functions on  $\Omega$  and is the pre-adjoint of the forward operator acting on the measures on  $\Omega$ . Since the elements  $w_i$  that figure in our catalogue are continuous functions, the distinction between  $A_0^*(E)$  and the Kolmogorov backward operator is inessential.

In words (2.29) states that the  $E$ -independent subspace spanned by  $\{w_1, w_2, \dots, w_k\}$  is in the domain of  $A_0^*(E)$  and invariant under  $A_0^*(E)$  for all relevant  $E$ . To avoid redundancy one should choose  $k$  as small as possible and we therefore require that the functions  $w_i : \Omega \rightarrow \mathbb{R}$ ,  $i = 1, 2, \dots, k$ , are linearly independent.

When (2.18) represents (2.19), condition (2.23) with  $P$  given by (2.28), amounts to

$$g(x, E)w_i'(x) - \mu(x, E)w_i(x) = \sum_{j=1}^k H_{ij}(E)w_j(x), \quad i = 1, 2, \dots, k. \quad (2.30)$$



It is easy to find a solution: if

$$\mu(x, E) = \gamma_0(x)g(x, E) + \mu_0(E) \quad (2.31)$$

for some functions  $\gamma_0 : \Omega \rightarrow \mathbb{R}$  and  $\mu_0 : \mathcal{E} \rightarrow \mathbb{R}$ , then the choice

$$\begin{aligned} k &= 1, \\ H(E) &= -\mu_0(E), \\ w(x) &= \exp\left(\int_{x_b}^x \gamma_0(y)dy\right), \text{ for some } x_b \in \Omega, \end{aligned} \quad (2.32)$$

makes (2.30) a valid equality. If we restrict to  $k = 1$ , then this is in fact the only possibility.

As mentioned above, condition (2.24) is a restriction on reproduction. The smaller the value of  $k$ , the more severe is the restriction. We therefore want to make  $k$  as large as possible while retaining the linear independence of  $\{w_1, w_2, \dots, w_k\}$ . So the question arises whether it is possible to pinpoint more restrictive conditions on  $g$  and  $\mu$  that allow for arbitrarily large values of  $k$ .

If the growth rate of an individual does not depend on its i-state but only on the environmental condition, that is, if

$$g(x, E) = v(E) \quad (2.33)$$

for some function  $v : \Omega \rightarrow \mathbb{R}$ , we call the i-state *physiological age* and often talk about maturation rather than growth. If on top of (2.33) we assume (2.31) and replace the unknown  $w$  by  $\tilde{w}$  via the transformation

$$w_i(x) = \exp\left(\int_{x_b}^x \gamma_0(y)dy\right) \tilde{w}_i(x), \quad i = 1, 2, \dots, k, \quad (2.34)$$

then (2.30) is equivalent to

$$\tilde{w}'(x) = \frac{\mu_0(E)}{v(E)} \tilde{w}(x) + \frac{1}{v(E)} H(E) \tilde{w}(x). \quad (2.35)$$

If we choose

$$H(E) = v(E)\Lambda - \mu_0(E)I, \quad (2.36)$$

with  $\Lambda$  an arbitrary  $k \times k$  matrix for arbitrary  $k$ , then the  $E$ -dependence vanishes from Eq. (2.35), which becomes an autonomous linear ODE with solution

$$\tilde{w}(x) = e^{(x-x_b)\Lambda} \tilde{w}(x_b). \quad (2.37)$$

To avoid redundancy, we have to make sure that  $\Lambda$  and  $\tilde{w}(x_b)$  are such that the components of  $\tilde{w}$  are linearly independent as scalar functions of the variable  $x \in \Omega$ , see conditions (H1) to (H3) of Sect. 5 as well as Corollary 6.3.

If instead of (2.33) we assume

$$g(x, E) = v_0(x)v_1(E) \quad (2.38)$$

for some functions  $v_0 : \Omega \rightarrow \mathbb{R}$  and  $v_1 : \mathcal{E} \rightarrow \mathbb{R}$ , we can introduce  $\zeta$  defined in terms of  $x$  by

$$\zeta(x) = \exp\left(\int_{x_b}^x \frac{dy}{v_0(y)}\right) \quad (2.39)$$

as a new i-state variable and thus reduce the situation to (2.33).

In Example 1.2 we have

$$g(x, E) = v_1(E) + v_2(E)x. \quad (2.40)$$

In combination with (2.31) this leads to

$$(v_1(E) + v_2(E)x) \tilde{w}'(x) = H_0(E) \tilde{w}(x), \quad (2.41)$$

where  $\tilde{w}$  is once more defined by (2.34) and where we have put

$$H_0(E) = H(E) + \mu_0(E)I. \quad (2.42)$$

For arbitrary  $k$  we can make (2.41) into an identity by choosing for  $i = 1, 2, \dots, k$

$$\tilde{w}_i(x) = x^{i-1} \quad (2.43)$$

and the entries of the matrix  $H_0(E)$

$$(H_0(E))_{ij} = \begin{cases} jv_1(E) & \text{if } j = i - 1, \\ (j - 1)v_2(E) & \text{if } j = i, \\ 0 & \text{otherwise.} \end{cases} \quad (2.44)$$

Again we can allow  $g$  to have an extra factor  $v_0(x)$  since we can remove it by the transformation from  $x$  to  $\zeta$  defined by (2.39).

If instead of (2.40) we assume that

$$g(x, E) = v_1(E) + v_2(E)x + v_3(E)x^2 \quad (2.45)$$

we can, as somewhat more complicated computations show, keep the  $\tilde{w}$  specified in (2.43), but adapt  $H_0(E)$  slightly as follows:

$$(H_0(E))_{ij} = -(k - i)v_3(E) \text{ if } j = i + 1, \quad (2.46)$$

while keeping the entries for all other combinations of  $i$  and  $j$  as in (2.44). But in addition we need to replace (2.31) by

$$\mu(x, E) = \gamma_0(x)g(x, E) + \mu_0(E) + (k - 1)v_3(E)x, \quad (2.47)$$

which, *nota bene*, involves  $k$ . So if we consider  $g$  and  $\mu$  as given, there can be at most one  $k$  for which this works. Again we can allow a factor  $v_0(x)$  in  $g$  and work with  $\zeta$  defined by (2.39).

The main result of the present paper is that problem (2.30), with  $\{w_1, w_2, \dots, w_k\}$  linearly independent, admits no other solution than the ones described above.

### 3 Physiologically structured population models

The formulation of a population model starts at the individual level with the specification of the *individual states* (i-states for short) representing physiological conditions that distinguish individuals from one another. The set of all admissible i-states is denoted by  $\Omega$ . In the present paper we restrict ourselves to the case of a finite dimensional  $\Omega \subset \mathbb{R}^n$  which we make a measurable space by equipping it with the  $\sigma$ -algebra  $\Sigma$  of all Borel sets. This measurable space is called the *individual state space* (i-state space). Our main results concern  $n = 1$  with  $\Omega$  an interval, possibly of infinite length. In that case one may think of the i-state as, for example, the size of an individual and we shall indeed often refer to the i-state as size.

The world in which individuals lead their lives has an impact on their development and behaviour. We capture the relevant aspects of the external world in a variable called the *environmental condition* denoted by  $E$  and taken from a set denoted by  $\mathcal{E}$ . One may think of  $E$  as a specification of food concentration, predation pressure and, possibly, other quantities like temperature or humidity.

Dependence among individuals arises from a feedback loop: the individuals themselves exert an influence on the environmental condition, for instance, by consuming food or serving as food for predators. As a rule, this feedback loop involves several species. We refer to the paper by Diekmann et al. (2010) for a concrete example. Note, however, that the example of cannibalism shows that this rule is not universal.

We consider the environmental condition as *input* and investigate how the input leads to *output* that comprises the contribution to the (change of) the environmental condition of the species itself, or other species, or any other quantity that we happen to be interested in. By taking population outputs as inputs for other populations or inanimate resources, we can build a dynamical model of a community. The ultimate model incorporates dependence among individuals and leads to nonlinear equations. But each building block considers the environmental condition as a given input and computes population output by summing the contributions by individuals. The present paper focusses on the *population state* (p-state for short) linear (but otherwise nonlinear) input–output map generated by a single building block.

The processes that have to be modelled are:

- i-state development (called *growth* for short),
- survival,

- reproduction (how much offspring and with what i-state at birth).

We assume that, given the environmental condition, growth is deterministic. We further assume that reproduction can be described by a per capita rate. We thus exclude, for instance, cell fission occurring exactly when the mother cell reaches a threshold size (see Example 8.5). Accordingly, and in line with Metz and Diekmann (1986) and de Roos et al. (2013), we introduce the three key model ingredients:

- the *growth rate*  $g(x, E)$ ,
- the *death rate*  $\mu(x, E)$ ,
- the *reproduction rate*  $\beta(x, E, \omega)$ ,  $\omega \in \Sigma$ .

Again we warn the readers that our use of the symbol  $\omega$  for a measurable subset of  $\Omega$  differs from the notational convention in probability theory.

The reproduction rate  $\beta$  should be interpreted as follows: the rate at which an individual of size  $x$  gives birth under environmental condition  $E$  is  $\beta(x, E, \Omega)$  and the state-at-birth of the offspring is distributed according to the Borel probability measure  $\beta(x, E, \cdot)/\beta(x, E, \Omega)$ .

Once we have a model at the i-level, it is a matter of bookkeeping to lift it to the p-level (Metz and Diekmann 1986; Diekmann and Metz 2010): Equating a p-level fraction to an i-level probability one obtains the deterministic (that is, the large population limit) link between the two levels. Still there are choices to be made for the formalism to employ: it could be partial differential equations (Metz and Diekmann 1986; Perthame 2007; Gwiazda and Marciniak-Czochra 2010) or renewal equations (RE) (Diekmann et al. 1998, 2001). Here we choose RE, albeit not in the most general form, since reproduction is described by a per capita rate.

As a first step we build composite ingredients from the basic ingredients  $g$ ,  $\mu$ , and  $\beta$ . We shall do so without specifying the nature of the environmental condition  $E$ , in particular, without specifying the space  $\mathcal{E}$  to which  $E$  belongs. Often we conceive of the environmental condition as a given function of time. When  $E$  occurs as a subscript to a function with argument  $(t, s)$  this entails that  $E(\tau)$  is given for  $\tau \in [s, t]$ .

We shall provide a constructive definition of the following quantities:

$X_E(t, s, x) :=$  the i-state at time  $t$ , given survival and given the input  
 $\tau \mapsto E(\tau)$ , of an individual that had i-state  $x \in \Omega$  at time  
 $s \leq t$ ,

$\mathcal{F}_E(t, s, x) :=$  the survival probability up to time  $t$ , given the input  $\tau \mapsto E(\tau)$ ,  
 of an individual that had i-state  $x \in \Omega$  at time  
 $s \leq t$ .

We assume that  $g$  and  $\tau \mapsto E(\tau)$  are such that the initial value problem

$$\frac{d\xi}{d\tau}(\tau) = g(\xi(\tau), E(\tau)), \quad \xi(s) = x \quad (3.1)$$

has a unique solution  $\xi(\tau)$  on  $[s, t]$  and define

$$X_E(t, s, x) := \xi(t). \quad (3.2)$$

We further *assume* that  $\mu$ ,  $g$ , and  $\tau \mapsto E(\tau)$  are such that also the initial value problem

$$\frac{df}{d\tau}(\tau) = -\mu(\xi(\tau), E(\tau))f(\tau), \quad f(s) = 1 \quad (3.3)$$

has a unique solution  $f(\tau)$  on  $[s, t]$  and define

$$\mathcal{F}_E(t, s, x) := f(t). \quad (3.4)$$

Let

$u_E^0(t, s, x, \omega) :=$  probability that, given the input  $\tau \mapsto E(\tau)$ ,  
an individual that had i-state  
 $x \in \Omega$  at time  $s \leq t$  is still alive at time  
 $t$  and has i-state in the set  $\omega \in \Sigma$ .

Then, since the growth of an individual is deterministic, we have

$$u_E^0(t, s, x, \omega) = \mathcal{F}_E(t, s, x) \delta_{X_E(t, s, x)}(\omega), \quad (3.5)$$

that is, the (unless survival is guaranteed) defective probability distribution  $u_E^0(t, s, x, \cdot)$  is a point measure concentrated at position  $X_E(t, s, x) \in \Omega$  of mass  $\mathcal{F}_E(t, s, x)$ . Let

$\beta_E^1(t, s, x, \omega) :=$  the rate at which, given the input  $\tau \mapsto E(\tau)$ ,  
an individual that had i-state  
 $x \in \Omega$  at time  $s \leq t$  produces at time  $t$   
offspring with i-state-at-birth in the set  $\omega \in \Sigma$ .

Then

$$\beta_E^1(t, s, x, \omega) = \mathcal{F}_E(t, s, x) \beta(X_E(t, s, x), E(t), \omega). \quad (3.6)$$

**Lemma 3.1** *Given the input  $\tau \mapsto E(\tau)$  defined on  $[s, t]$ , the relations*

$$u_E^0(t, s, x, \omega) = \int_{\Omega} u_E^0(t, \tau, y, \omega) u_E^0(\tau, s, x, dy), \quad (3.7)$$

$$\begin{aligned} \beta_E^1(t, s, x, \omega) &= \int_{\Omega} \beta_E^1(t, \tau, y, \omega) u_E^0(\tau, s, x, dy) \\ &= \mathcal{F}_E(\tau, s, x) \beta_E^1(t, \tau, X_E(\tau, s, x), \omega) \end{aligned} \quad (3.8)$$

hold for all  $\tau \in (s, t)$  and all  $\omega \in \Sigma$ .

Relation (3.7) is the Chapman–Kolmogorov equation and (3.8) is a similar consistency relation relating growth, survival and reproduction.

We omit the straightforward proof of Lemma 3.1, but note that, essentially, the consistency relations reflect the uniqueness of solutions to Eqs. (3.1) and (3.3).

The composite ingredients  $u_E^0$  and  $\beta_E^1$  satisfying (3.7) and (3.8) are the starting point for a next round of constructive definitions. They are examples of *kernels parametrised by the input* (cf. Diekmann et al. 2001). Such a kernel  $\phi_E$  assigns to each input  $\tau \mapsto E(\tau)$  defined on  $[s, t]$  a function  $\phi_E(t, s, \cdot, \cdot) : \Omega \times \Sigma \rightarrow \mathbb{R}$  which is bounded and measurable with respect to the first variable and countably additive with respect to the second variable. This is to say that for fixed  $\omega \in \Sigma$  the function  $x \mapsto \phi_E(t, s, x, \omega)$  is bounded and measurable while for fixed  $x \in \Omega$  the map  $\omega \mapsto \phi_E(t, s, x, \omega)$  is a finite signed measure.

The product  $(\phi \star \psi)_E$  of two kernels  $\phi_E$  and  $\psi_E$  parametrised by the input is defined by

$$(\phi \star \psi)_E(t, s, x, \omega) = \int_s^t \int_{\Omega} \phi_E(t, \tau, y, \omega) \psi_E(\tau, s, x, dy) d\tau. \quad (3.9)$$

The  $\star$ -product is associative.

For  $k \geq 2$  we define recursively

$$\beta_E^k := \left( \beta^1 \star \beta^{k-1} \right)_E \quad (3.10)$$

The interpretation of  $\beta_E^k$  is as follows: Given the input  $\tau \mapsto E(\tau)$  defined on  $[s, t]$ , the quantity  $\beta_E^2(t, s, x, \omega)$  is the rate at which grandchildren to an individual that had i-state  $x$  at time  $s$  are born at time  $t$  with i-state-at-birth in the set  $\omega \in \Sigma$ . The quantity  $\beta_E^3(t, s, x, \omega)$  has the same interpretation but for great-grandchildren and  $\beta_E^k(t, s, x, \omega)$  for  $k$ th generation offspring. To get the combined birth rate of all descendants of such an individual we sum up over all generations and define

$$\beta_E^c := \sum_{k=1}^{\infty} \beta_E^k, \quad (3.11)$$

where the superscript  $c$  stands for *clan*.

Because every member of the clan is either a child of the ancestor or a child of a member of the clan, or, alternatively, either a child of the ancestor or a member of the clan of a child of the ancestor, we obtain a consistency relation in terms of the following RE:

$$\beta_E^c = \beta_E^1 + \left( \beta^1 \star \beta^c \right)_E = \beta_E^1 + \left( \beta^c \star \beta^1 \right)_E. \quad (3.12)$$

Mathematically, (3.12) means that  $\beta_E^c$  is the *resolvent kernel* of the kernel  $\beta_E^1$  (cf. Gripenberg et al. 1990).

In order to incorporate both the founding ancestor and the development of the descendants after birth, we finally define

$$u_E^c = u_E^0 + \left(u^0 \star \beta^c\right)_E. \quad (3.13)$$

For later use we note that

$$\beta_E^c(s, s, x, \omega) = \beta_E^1(s, s, x, \omega) = \beta(x, E(s), \omega) \quad (3.14)$$

and

$$u_E^c(s, s, x, \omega) = u_E^0(s, s, x, \omega) = \delta_x(\omega). \quad (3.15)$$

The clan-kernels  $u_E^c$  and  $\beta_E^c$  satisfy the same consistency relations as  $u_E^0$  and  $\beta_E^1$  of Lemma 3.1.

**Theorem 3.2** *Given the input  $\tau \mapsto E(\tau)$  defined on  $[s, t]$ , the relations*

$$u_E^c(t, s, x, \omega) = \int_{\Omega} u_E^c(t, \tau, y, \omega) u_E^c(\tau, s, x, dy), \quad (3.16)$$

$$\beta_E^c(t, s, x, \omega) = \int_{\Omega} \beta_E^c(t, \tau, y, \omega) u_E^c(\tau, s, x, dy) \quad (3.17)$$

hold for all  $\tau \in (s, t)$  and all  $\omega \in \Sigma$ .

The proof of Theorem 3.2 proceeds by proving (3.17) first and next use this identity to verify (3.16). The papers (Diekmann et al. 1998, 2001, 2003) contain a much more detailed exposition of this constructive approach, including proofs of (3.16) and (3.17) in a generalised form with the instantaneous rate  $\beta_E^1$  replaced by cumulative offspring production  $\Lambda$ , necessitating the use of the Stieltjes integral, which can be avoided here since we assume that offspring are produced at a per capita rate, so with some probability per unit of time. In our paper (Diekmann et al. 1998) we considered general linear time dependent problems (the time dependence corresponding to fixing an input). In the paper (Diekmann et al. 2001) we explicitly considered input, but focussed on the feedback loop that captures dependence. As the notation of (Diekmann et al. 2001) is not ideal for investigating the problem introduced in the next section, we have adopted in the present exposition a different, more suitable notation, viz. the use of the subscript  $E$ . In the paper (Diekmann et al. 2003) the main objective was to characterise the steady states.

The central idea of the modelling and analysis methodology of physiologically structured populations is that we view the population state as a measure  $m$  on  $\Omega$ . We use the kernel  $u_E^c$  to define operators  $U_E^c(t, s)$  that map the p-state at time  $s$  to the p-state at time  $t$  as follows. Assuming that all individuals experience the same environmental condition  $E$ , we can associate to each measure  $m$  a new measure

$$(U_E^c(t, s)m)(\omega) := \int_{\Omega} u_E^c(t, s, x, \omega) m(dx) \quad (3.18)$$

and note that the Chapman–Kolmogorov relation (3.16) translates into the semigroup property

$$U_E^c(t, s) = U_E^c(t, \tau)U_E^c(\tau, s), \quad s < \tau < t, \quad (3.19)$$

while (3.15) yields

$$U_E^c(s, s) = I. \quad (3.20)$$

Note that we may replace the superscript  $c$  in (3.18) by 0 and use (3.7) to deduce the semigroup property for  $U_E^0$ . That  $U_E^0(s, s) = I$  follows again from (3.15). Families of linear maps that satisfy the conditions (3.18) and (3.20) are called *state-linear dynamical systems with input*.

In a previous paper (Diekmann et al. 2018) we have already considered what amounts to population level conditions for ODE-reducibility, in this paper we concentrate on finding i-level ones.

## 4 Finite dimensional state representation

### 4.1 General considerations

Let  $Y$  be a vector space. We do not yet fix a topology for  $Y$ , but note that if  $Y'$  is a separating vector space of linear functionals on  $Y$ , then the weakest topology on  $Y$  for which all  $y' \in Y'$  are continuous (the so-called  $w(Y, Y')$ -topology) makes  $Y$  into a locally convex space whose dual space is  $Y'$  (Rudin 1973; Theorem 3.10, p.62).

Let  $U_E$  be a state-linear dynamical system with input  $\tau \mapsto E(\tau)$ , which at this point has no connection yet with either  $U_E^0$  or  $U_E^c$ . This means that each  $U_E(t, s)$  is a linear operator on  $Y$  and

$$U_E(s, s) = I, \quad (4.1)$$

$$U_E(t, s) = U_E(t, \tau)U_E(\tau, s), \quad s < \tau < t. \quad (4.2)$$

If a vector topology has been chosen for  $Y$  we also require the operators  $U_E(t, s)$  to be continuous with respect to this topology.

We are interested in finding a finite dimensional exact reduction (or, lumping) of  $U_E(t, s)$ . More precisely, we want, if possible, to choose a separating vector space  $Y'$  of linear functionals on  $Y$  and construct a  $w(Y, Y')$ -continuous linear map  $P : Y \rightarrow \mathbb{R}^k$  and a  $k \times k$  matrix  $K(E)$  such that

$$PU_E(t, s) = \Phi_E(t, s)P, \quad (4.3)$$

where  $\Phi_E(t, s)$  is the fundamental matrix solution of the  $k$ -dimensional ODE-system

$$\frac{dN}{d\tau}(\tau) = K(E(\tau))N(\tau). \quad (4.4)$$



The  $w(Y, Y')$ -continuous linear map  $P$  can be represented by elements  $W_i \in Y'$ ,  $i = 1, 2, \dots, k$  as

$$(Py)_i = \langle y, W_i \rangle. \quad (4.5)$$

To avoid useless variables,  $P$  should be surjective or, equivalently, the functionals  $W_1, W_2, \dots, W_k$  should be linearly independent.

The adjoint of a forward evolutionary system characterised by (4.1)–(4.2) is a backward evolutionary system (see Clément et al. 1988; Diekmann et al. 1995). For these it is more natural to think of  $s$  as the dynamic variable with respect to which we differentiate. Since the restriction of an evolutionary system to the diagonal in the  $(t, s)$ -plane is the identity operator, the derivative with respect to  $t$  is simply minus the derivative with respect to  $s$  at the diagonal. This observation suffices for our purpose and we therefore do not elaborate the forward-backward duality here.

Let  $W$  denote the  $k$ -vector with components  $W_i$ . We may then rewrite (4.3) in the form

$$(U_E(t, s))^* W = \Phi_E(t, s) W \quad (4.6)$$

as a shorthand for

$$(U_E(t, s))^* W_i = \sum_{j=1}^k (\Phi_E(t, s))_{ij} W_j, \quad i = 1, 2, \dots, k. \quad (4.7)$$

Since the right hand side of (4.7) is differentiable, the same must be true for the left hand side. By differentiation we obtain the following task:

$$\text{TASK}_L: \quad \frac{d}{dt} (U_E(t, s))^* W \Big|_{t=s} = K(E(s)) W.$$

Here the subscript L refers to *lumpability* and the task is to find  $k$  elements  $W_i \in Y^*$  such that

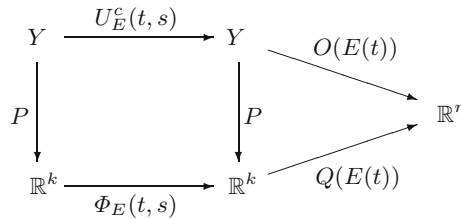
- the derivatives exist,
- the outcome is a linear combination, with input dependent coefficients, of the elements  $W_i$ .

Assuming that the derivative exists, we may write

$$\frac{d}{dt} (U_E(t, s))^* W \Big|_{t=s} = A(E(s))^* W \quad (4.8)$$

and in Sects. 2.1 and 8 we employ the notation of the right hand side.

In this paper we accomplish  $\text{TASK}_L$  for the state-linear dynamical system  $U_E^0$  with input introduced in Sect. 3. In fact, we characterise the growth rates  $g$  and the death rates  $\mu$  for which exact reduction is possible and we compute the corresponding  $W$



**Fig. 2** Structure of models with output and finite dimensional reduction

and  $K(E)$ . In order to explain the relevance of these results for the dynamical system  $U_E^c$  we now widen the perspective by introducing output.

We return to (4.1) and (4.2) and complement them by a  $w(Y, Y')$ -continuous linear map

$$O(E) : Y \rightarrow \mathbb{R}^r. \quad (4.9)$$

We call

$$O(E(t))U_E(t, s)y \quad (4.10)$$

the *output* at time  $t$ , given the state  $y$  at time  $s$  and the input  $\tau \mapsto E(\tau)$  defined on  $[s, t]$ . The idea is that the state itself is not observable, only the output can be measured. We now ask the following question: can, in fact, the relation between the pair (initial state, input) on the one hand and output on the other, alternatively and equivalently be described in terms of a *finite dimensional* dynamical system? That is, when is the diagram in Fig. 2 commutative for all inputs  $\tau \mapsto E(\tau)$  defined on  $[s, t]$ ?

More precisely, we again want to find an integer  $k$  and continuous linear maps  $P$  and  $K(E)$ , but in addition to (4.3) we now require

$$O(E(t))U_E(t, s) = Q(E(t))\Phi_E(t, s)P, \quad (4.11)$$

where the  $k \times r$  matrix  $Q(E)$  is also to be determined. In other words, we require that, given the state  $y$  at time  $s$  and the input  $\tau \mapsto E(\tau)$  defined on  $[s, t]$ , the output at time  $t$  is obtained by applying  $Q(E(t))$  to the solution at time  $t$  of (4.4) with initial condition

$$N(s) = Py. \quad (4.12)$$

We stress that, as far as the output is concerned, the reduction does *not* involve loss of information. At the black box level we cannot distinguish between the true system  $U_E$  and its finite dimensional counterpart  $\Phi_E$ .

**Remark 4.1** In principle we could allow  $P$  to depend on  $E$  and write the initial condition (4.12) as

$$N(s) = P(E(s))y. \quad (4.13)$$

But if we can choose  $E(s)$  without affecting  $U_E(t, s)$ , then, since also  $\Phi_E(t, s)$  is insensitive to the precise value of  $E$  in  $s$ , it follows that after all,  $P$  cannot depend on  $E$ . A similar observation was made for a related problem in Remark 7.3 of (Diekmann et al. 2018).

Taking  $t = s$  in (4.11) we find that necessarily

$$O(E) = Q(E)P. \quad (4.14)$$

This allows us to rewrite (4.11) as

$$Q(E(t)) (PU_E(t, s) - \Phi_E(t, s)P) = 0. \quad (4.15)$$

The aim is now to derive necessary and sufficient conditions on  $O(E)$  and  $U_E$  for the existence of  $P$ ,  $Q(E)$ ,  $K(E)$  that make (4.15) a valid identity for a *minimal value* of  $k$ . Note that for  $k > r$  the Eq. (4.15) allows for possibly redundant information in (4.4). Indeed, adding components to  $N$  that are unobservable, in the sense that they do not contribute directly or indirectly to future output, does not harm; by requiring  $k$  to be minimal we avoid such redundancy. To derive the conditions, we follow an iterative procedure that is well-known in systems theory.

Our starting point is the output map  $O(E)$ . For (4.14) to be possible at all, the range of  $O(E)^* : \mathbb{R}^r \rightarrow Y'$  should be contained in a finite dimensional subspace which is independent of  $E$ . Without loss of generality we can assume that this subspace has dimension  $r$ . Indeed, if the dimension is less than  $r$  there is dependence among the output components and by choosing suitable coordinates we can reduce  $r$  without losing any information.

Let  $\{W_1^0, W_2^0, \dots, W_r^0\} \subset Y'$  be a basis for the subspace of  $Y'$  that contains the range of  $O(E)^*$ . Recall the representation (4.5) of  $P$  in terms of the elements  $W_i \in Y'$ ,  $i = 1, 2, \dots, k$ . From (4.14) we conclude that for all  $i = 1, 2, \dots, r$ , the element  $W_i^0$  belongs to the subspace spanned by  $\{W_1, W_2, \dots, W_k\}$ . In particular,  $k \geq r$ . By a suitable choice of basis for  $\mathbb{R}^k$  we can arrange things so that

$$W_i = W_i^0, \quad i = 1, 2, \dots, r. \quad (4.16)$$

Define  $P_0 : Y \rightarrow \mathbb{R}^r$  by

$$(P_0 y)_i = \langle y, W_i^0 \rangle, \quad i = 1, 2, \dots, r. \quad (4.17)$$

Then

$$O(E) = Q_0(E)P_0 \quad (4.18)$$

with  $Q_0(E) : \mathbb{R}^r \rightarrow \mathbb{R}^r$  such that for all  $v \in \mathbb{R}^r$

$$O(E)v = 0 \text{ for all } E \in \mathcal{E} \Rightarrow v = 0 \quad (4.19)$$

because we have chosen  $r$  in the optimal way and  $\{W_1^0, W_2^0, \dots, W_r^0\}$  is a basis for the range of  $O(E)^*$ . In the first step in the iterative procedure we try to find  $K_0(E) : \mathbb{R}^r \rightarrow \mathbb{R}^r$  such that (4.15) in the guise of

$$Q_0(E(t)) (P_0 U_E(t, s) - \Phi_E(t, s) P_0) = 0. \quad (4.20)$$

holds, where  $\Phi_E(t, s)$  is now the fundamental matrix solution of the  $r$ -dimensional system

$$\frac{dN}{d\tau}(\tau) = K_0(E(\tau))N(\tau). \quad (4.21)$$

On account of (4.19) we may reduce (4.20) to

$$P_0 U_E(t, s) = \Phi_E(t, s) P_0 \quad (4.22)$$

if, as we assume, we can manipulate  $E(t)$  without changing  $P_0 U_E(t, s)$ .

One should compare (4.22) with (4.3), but there is an important difference: determining the map  $P$  or, equivalently, the elements  $W_i$ ,  $i = 1, 2, \dots, k$  of  $Y'$  is part of the task whereas the elements  $W_i^0$ ,  $i = 1, 2, \dots, r$  are known because  $O(E)$  is given. So rather than a task, we now have the following test:

$$\text{TEST}_0 : \quad \frac{d}{dt}(U_E(t, s))^* W^0 \Big|_{t=s} \stackrel{?}{=} K_0(E(s)) W^0.$$

In more detail the test consists in answering the following questions:

- Does  $\frac{d}{dt} U_E^*(t, s) W_i^0 \Big|_{t=s}$  exist for  $i = 1, 2, \dots, r$ ?
- Is the outcome a linear combination, with input dependent coefficients, of the elements  $W_1^0, W_2^0, \dots, W_r^0$ ?

If, for some index  $i$ , the derivative does not exist, finite dimensional state representation is not possible. In contrast, there is still hope that a finite dimensional state representation might be possible if a derivative is not in the span of  $\{W_1^0, W_2^0, \dots, W_r^0\}$ . If that is the case, we add the derivative to the basis and thus enlarge the subspace. Varying both the index  $i$  and the value of  $E(s)$ , we obtain a new subspace of  $Y'$  that may, or may not, be finite dimensional. If it is finite dimensional we perform  $\text{TEST}_1$  which is the analogue of  $\text{TEST}_0$ , but with  $W^0$  extended to  $W^1$ , a vector the components of which span the new subspace. If necessary this procedure can be repeated. If the process leads after a finite number of steps to a finite dimensional subspace, we are in business. If the process does not terminate, finite dimensional state representation is not possible.

In general, finding  $P$  and  $K(E)$  such that (4.3) holds, or, in other words, performing  $\text{TASK}_L$ , is hard since, literally, one does not know how to start. In contrast, for a given output  $O(E)$ , the tests  $\text{TEST}_0, \text{TEST}_1, \dots$  yield a constructive procedure.

If one does manage to characterise  $P$  and  $K(E)$  such that (4.3) holds, one can give the output problem a twist: (4.11) holds for all outputs of the form (4.14). Below we

shall follow this road while considering reproduction as part of the output. In this manner we can focus on (4.3) for  $U_E^0$  and in the end still obtain results for  $U_E^c$  as described in the next subsection.

## 4.2 Physiologically structured population models

As explained in Sect. 3, it is natural to consider the p-state of a physiologically structured population as a measure and therefore the p-state space should be a linear subspace  $Y$  of the vector space  $M(\Omega)$  of all Borel measures on  $\Omega$ . One reason for not choosing the whole  $M(\Omega)$  as p-state space is that we may want to keep biologically relevant quantities such as the total biomass finite. If in the one-dimensional i-state space case  $x$  denotes the size of an individual, then  $\int_{\Omega} x m(dx)$  represents the total biomass and it is finite only for measures  $m$  in a proper subspace of  $M(\Omega)$  if  $\Omega$  is an infinite interval. An other reason is that when we check whether a reduction of the infinite dimensional model is ODE-reducible we construct a  $w(Y, Y')$ -continuous linear map  $P : Y \rightarrow \mathbb{R}^k$  or, equivalently, linear functionals  $W_i \in Y'$ , now representable by locally bounded measurable functions  $w_i$  on  $\Omega$  via the pairing

$$(Pm)_i = \langle m, W_i \rangle = \int_{\Omega} w_i(x) m(dx), \quad (4.23)$$

and we may end up with functions  $w_i$  for which the integral in (4.23) is not finite for every  $m \in M(\Omega)$ . So we want these functions  $w_i$  to represent elements in  $Y'$  and consequently have to restrict  $Y$  to a suitably chosen subspace of  $M(\Omega)$ . The freedom we have in choosing  $Y$  and  $Y'$  therefore comes in very handy.

We denote the  $\mathbb{R}^k$ -valued function with components  $w_i$  by  $w$ . Later we shall show that when a finite dimensional state representation exists for the physiologically structured population model, the function  $w$  is actually continuous, but not necessarily bounded, on  $\Omega$ .

Consider the dynamical systems  $U_E^0$  and  $U_E^c$  with input of Sect. 3. The system  $U_E^0$  represents a transport-degradation process (without reproduction) while  $U_E^c$  represents a physiologically structured population model with reproduction.

Using (3.18), in its superscript 0 version, (3.13) and (3.5) we find that for the transport-degradation case, (4.3) amounts to

$$\mathcal{F}_E(t, s, x) w(X_E(t, s, x)) = \Phi_E(t, s) w(x), \quad (4.24)$$

while for complete physiologically structured population models, it amounts to

$$\begin{aligned} & \mathcal{F}_E(t, s, x) w(X_E(t, s, x)) \\ & + \int_s^t \int_{\Omega} \mathcal{F}_E(t, \tau, y) w(X_E(t, \tau, y)) \mathcal{F}_E(\tau, s, x) \beta_E^c(\tau, s, x, dy) d\tau \\ & = \Phi_E(t, s) w(x). \end{aligned} \quad (4.25)$$

Using in addition (3.14) and (3.5) we can elaborate  $\text{TASK}_L$  by taking the derivative with respect to  $t$  of both sides of (4.25) and then putting  $t = s$ . Since  $t \mapsto X_E(t, s, x)$  is differentiable, the differentiability of the left hand side of (4.25) implies that the function  $x \mapsto w(x)$  is differentiable, at least in certain directions. More precisely, for  $x \in \Omega$  the linear approximation of  $w(x + h) - w(x)$  exists as a map acting on the space of vectors  $h$  spanned by  $g(x, \cdot)$ . This map we call  $(Dw)(x)$ . We therefore find that necessarily

$$(Dw)(x)g(x, E) - \mu(x, E)w(x) = K(E)w(x). \quad (4.26)$$

for the transport-degradation model, and

$$(Dw)(x)g(x, E) - \mu(x, E)w(x) + \int_{\Omega} w(y)\beta(x, E, dy) = K(E)w(x). \quad (4.27)$$

for the population model.

If the transport-degradation model is ODE-reducible, that is, if we can find a positive integer  $k$ , linearly independent measurable functions  $\{w_1, w_2, \dots, w_k\}$  and a  $k \times k$  matrix  $K(E)$  so that (4.26) holds, the physiologically structured population model is also ODE-reducible if we impose the following restriction on the reproduction process: There exists a  $k \times k$  matrix  $M(E)$  such that

$$\int_{\Omega} w(y)\beta(x, E, dy) = M(E)w(x). \quad (4.28)$$

The reason is simply that when this is the case, we can write the Eqs. (4.26) and (4.27) in a unified way as

$$(Dw)(x)g(x, E) = (H(E) + \mu(x, E)I)w(x), \quad (4.29)$$

where  $H(E) = K(E)$  for the transport-degradation model and  $H(E) = K(E) - M(E)$  for the population model.

Note that the restriction (4.28) is satisfied if reproduction is (part of the) output in the following sense:

$$\beta(x, E, \cdot) = \sum_{j=1}^k \beta_j(E, \cdot)w_j(x), \quad (4.30)$$

because then we can take  $M(E)$  to be the  $k \times k$  matrix with entries

$$M_{ij}(E) = \int_{\Omega} w_i(y)\beta_j(E, dy). \quad (4.31)$$

But there are other situations in which (4.28) is satisfied. The simplest case is when  $M(E) = 0$ . Note that this does not imply the absence of reproduction, merely that  $\beta$  is annihilated by all  $w_i$  or, in other words, that the functions  $w_i$  measure properties

of individuals that are preserved under reproduction (as a concrete example, think of mass in cell-fission models).

Yet another case, viz. the one in which  $\beta$  has the form

$$\beta(x, E, \omega) = \tilde{\beta}(x, E)\beta_0(\omega), \quad (4.32)$$

will be briefly discussed in Sect. 8.

When  $n = 1$ , that is, when the i-state space is one-dimensional, (4.29) reduces to the following equation:

$$g(x, E)w'(x) = (H(E) + \mu(x, E)I)w(x). \quad (4.33)$$

In Sect. 5 we address the following problem: list necessary and sufficient conditions on  $g$  and  $\mu$  for the existence of a measurable function  $w : \Omega \rightarrow \mathbb{R}^k$  with a priori unknown  $k$  such that there exists a  $k \times k$  matrix  $H(E)$  for which (4.33) holds.

To solve the problem, we heavily use that (4.33) is a local equation in the  $x$ -variable. For a fixed value  $E_0$  the solution  $w$  of (4.33) is, essentially, given by a matrix exponential. Once  $w$  has been determined, we can view (4.33) as a constraint for the ways in which  $g$ ,  $\mu$  and  $H$  can depend on  $E$ .

## 5 A catalogue of models that admit a finite dimensional state representation

In this section we present an explicit catalogue of all possible combinations of model ingredients that allow a finite dimensional state representation for transport-degradation models with  $\Omega \subset \mathbb{R}^1$  an interval that may have infinite length. The catalogue extends naturally to all physiologically structured population models in which the submodel for reproduction is restricted by (4.28).

One should choose the individual state space  $\Omega$  such that every point in it is reachable. The following assumption guarantees this. We shall also make use of it in the proofs of our results.

**Assumption 5.1** There exists a constant environmental condition  $E_0$  such that

$$g(x, E_0) > 0 \quad \text{for all } x \in \Omega.$$

The catalogue consists of three families  $F_i$ ,  $i = 1, 2, 3$ , of functions  $g$  and  $\mu$  for which we specify  $w$  and  $H(E)$  such that (4.33) holds. These families involve infinite dimensional parameters in the form of functions  $\gamma_0 : \Omega \rightarrow \mathbb{R}$ ,  $\mu_0 : \mathcal{E} \rightarrow \mathbb{R}$ ,  $v_0 : \Omega \rightarrow \mathbb{R}$ ,  $v_i : \mathcal{E} \rightarrow \mathbb{R}$ ,  $i = 1, 2, 3$ . We do not claim that  $g$ ,  $\mu$  and  $w$  are biologically meaningful for all choices of these parameters (in fact they are not). The catalogue simply provides a precise description of the constraints on  $g$  and  $\mu$  that enable an equivalent, as far as the output map defined by  $w$  is concerned, finite dimensional representation of the corresponding transport-degradation model.

A transformation of the i-state variable affects the death rate  $\mu$  and the output function  $w$  in the usual manner. But since the growth rate needs to keep its interpretation, we have to incorporate a transformation specific factor. If

$$y = \phi(x) \quad (5.1)$$

and

$$\frac{dx}{da} = g(x, E), \quad (5.2)$$

then

$$\begin{aligned} \frac{dy}{da} &= \phi'(x) \frac{dx}{da} = \phi'(x) g(x, E) = \phi'(\phi^{-1}(y)) g(\phi^{-1}(y), E) \\ &= \frac{1}{(\phi^{-1})'(y)} g(\phi^{-1}(y), E) \end{aligned} \quad (5.3)$$

and accordingly

$$\tilde{g}(y, E) = \frac{1}{(\phi^{-1})'(y)} g(\phi^{-1}(y), E) \quad (5.4)$$

is the growth rate of the transformed i-state variable  $y$ .

Transformation of the i-state-variable induces a transformation of the parameters in our families  $F_i$ . As we shall indicate below, the multiplicative factor highlighted by (5.4) allows us to transform the i-state variable in such a way that the formula for the growth rate becomes relatively simple.

Let  $L : \mathbb{R}^k \rightarrow \mathbb{R}^k$  be linear and invertible. If (4.33) holds and we define

$$\begin{aligned} \tilde{w} &= Lw, \\ \tilde{H} &= LHL^{-1}, \end{aligned} \quad (5.5)$$

then, by applying  $L$  to the identity (4.33), we obtain

$$g(x, E) \tilde{w}'(x) = (\tilde{H}(E) + \mu(x, E)I) \tilde{w}(x). \quad (5.6)$$

So, if  $g$  and  $\mu$  are given and  $(w, H)$  is a solution to (4.33), then  $(\tilde{w}, \tilde{H})$  defined by (5.5) is also a solution. As (5.5) defines an equivalence relation we see that solutions  $(w, H)$  to (4.33) occur in equivalence classes. In our catalogue we choose  $w$  such that  $H$  has a relatively simple form, but one should keep in mind that this choice yields a representative of an equivalence class.

As a reference for integration we choose a reference point  $x_b \in \Omega$ . If all individuals are born with the same i-state at birth, we choose this i-state as  $x_b$ .

We are now ready to present our first family in the catalogue.



---

 $F_1$ : scalar representations

---

$k$	$k = 1$
Parameters	$\gamma_0 : \Omega \rightarrow \mathbb{R}, \mu_0 : \mathcal{E} \rightarrow \mathbb{R}$
$g$	No restriction
$\mu$	$\mu(x, E) = \gamma_0(x)g(x, E) + \mu_0(E)$
$w$	$w(x) = \exp\left(\int_{x_b}^x \gamma_0(y)dy\right)$
$H$	$H(E) = -\mu_0(E)$

---

The proof that (4.33) is an identity if  $F_1$  applies is straightforward and omitted.

In the rest of this section we assume that  $k \geq 2$  and we supplement (4.33) with the non-degeneracy condition

(H1)  $w_1, w_2, \dots, w_k$  are linearly independent as scalar functions of  $x \in \Omega$ .

Among the parameters of the second family are a constant (that is, independent of both  $x$  and  $E$ )  $k \times k$  matrix  $\Lambda$  and a constant vector  $w(x_b) \in \mathbb{R}^k$  specifying the value of  $w$  in  $x_b$ . The identity (4.33) holds for all choices of  $\Lambda$  and  $w(x_b)$ . But (H1) has also to be satisfied. Therefore we have to impose the condition

(H2) the eigenvalues of  $\Lambda$  have geometric multiplicity one  
on the matrix  $\Lambda$  and the condition

(H3) if  $\{\theta_1, \theta_2, \dots, \theta_k\}$  is a basis for  $\mathbb{R}^k$  consisting of eigenvectors and generalised eigenvectors of  $\Lambda$  and

$$w(x_b) = \sum_{j=1}^k d_j \theta_j,$$

then  $d_j \neq 0$  whenever  $\theta_j$  is a generalised eigenvector of highest rank on the combination of  $\Lambda$  and  $w(x_b)$ .

We are now ready for the second family in our catalogue.

---

 $F_2$ : physiological age

---

$k$	$k \geq 2$
Parameters	$\gamma_0 : \Omega \rightarrow \mathbb{R}, \mu_0 : \mathcal{E} \rightarrow \mathbb{R}, v_0 : \Omega \rightarrow \mathbb{R}, v_1 : \mathcal{E} \rightarrow \mathbb{R}$ $\Lambda \in \mathbb{R}^{k \times k}$ and $w(x_b) \in \mathbb{R}^k$ such that (H2) and (H3) hold. $\zeta(x) := \int_{x_b}^x \frac{dy}{v_0(y)}$
$g$	$g(x, E) = v_0(x)v_1(E)$
$\mu$	$\mu(x, E) = \gamma_0(x)g(x, E) + \mu_0(E)$
$w$	$w(x) = \exp\left(\int_{x_b}^x \gamma_0(y)dy\right) \exp(\zeta(x)\Lambda) w(x_b)$
$H$	$H(E) = v_1(E)\Lambda - \mu_0(E)I$

---

Note that we do not lose any generality by assuming that  $\Lambda$  is in Jordan normal form. Apart from the common factor  $\exp\left(\int_{x_b}^x \gamma_0(y)dy\right)$  the components of  $w$  are therefore linear combinations of  $k$  building blocks of the form

$$(\zeta(x))^m e^{\lambda_j \zeta(x)}, \quad (5.7)$$

where  $j = 1, 2, \dots, r$ ,  $m = 0, 1, \dots, k_j - 1$ ,  $\sum_{j=1}^r k_j = k$ . The condition (H3) guarantees that each and every building block (5.7) contributes to at least one component of  $w$ . By choosing  $w(x_b)$  such that  $d_j = 1$  whenever  $\theta_j$  is a generalised eigenvector of highest rank and zero otherwise, the components are (because  $\Lambda$  is assumed to be in Jordan normal form) precisely these building blocks.

We now present the third and last family of the catalogue.

---

$F_3$ : generalised von Bertalanffy growth

---

$k$	$k \geq 2$
Parameters	$\gamma_0 : \Omega \rightarrow \mathbb{R}, \mu_0 : \mathcal{E} \rightarrow \mathbb{R}, v_0 : \Omega \rightarrow \mathbb{R}, v_j : \mathcal{E} \rightarrow \mathbb{R}, j = 1, 2, 3.$ $\zeta(x) := \int_{x_b}^x \frac{dy}{v_0(y)}$
$g$	$g(x, E) = v_0(x) \left( v_1(E) + v_2(E)\zeta(x) + v_3(E)\zeta(x)^2 \right)$
$\mu$	$\mu(x, E) = \gamma_0(x)g(x, E) + \mu_0(E) + (k-1)v_3(E)\zeta(x)$
$w$	$w_j(x) = \exp \left( \int_{x_b}^x \gamma_0(y) dy \right) \zeta(x)^{j-1}, \quad j = 1, 2, \dots, k.$
$H$	$H(E) = H_0(E) - \mu_0(E)I$ $H_0 = \begin{pmatrix} 0 & -(k-1)v_3 & 0 & \dots & 0 \\ v_1 & v_2 & -(k-2)v_3 & & 0 \\ 0 & 2v_1 & 2v_2 & & 0 \\ \vdots & 0 & 3v_1 & & 0 \\ & & & \ddots & \\ 0 & \dots & \dots & (k-1)v_1 & (k-1)v_2 \end{pmatrix}$

---

It is important to realise that the parametrisation is far from unique. For instance, if we write

$$g = \frac{1}{\zeta'} \left( v_1 + v_2 \zeta + v_3 \zeta^2 \right)$$

it is not difficult to check that the linear fractional transformation

$$\eta = \frac{a\zeta + b}{c\zeta + d}$$

with  $ad - bc \neq 0$  yields the alternative form

$$g = \frac{1}{\eta'} \left( \tilde{v}_1 + \tilde{v}_2 \eta + \tilde{v}_3 \eta^2 \right)$$

with

$$\tilde{v}_1 = \frac{a^2 v_1 - ab v_2 + b^2 v_3}{ad - bc},$$

$$\begin{aligned}\tilde{v}_2 &= \frac{-2acv_1 + (ad + bc)v_2 - 2bdv_3}{ad - bc}, \\ \tilde{v}_3 &= \frac{c^2v_1 - cdv_2 + d^2v_3}{ad - bc}.\end{aligned}$$

With a little bit more effort one can check that

$$\mu = \gamma_0 g + \mu_0 + (k-1)v_3\zeta$$

becomes

$$\mu = \tilde{\gamma}_0 g + \tilde{\mu}_0 + (k-1)\tilde{v}_3\eta$$

with

$$\begin{aligned}\tilde{\gamma}_0 &= \gamma_0 + \frac{c(k-1)\zeta'}{c\zeta + d}, \\ \tilde{\mu}_0 &= \mu_0 - (k-1)(acv_1 - bcv_2 + bdv_3).\end{aligned}$$

Note that

$$\int_{x_b}^x \tilde{\gamma}_0(y)dy = \int_{x_b}^x \gamma_0(y)dy + (k-1)\log\left(\frac{c\zeta + d}{d}\right)$$

and that, accordingly,

$$\begin{aligned}\tilde{w}_j(x) &= \exp\left(\int_{x_b}^x \tilde{\gamma}_0(y)dy\right) \eta(x)^{j-1} \\ &= \exp\left(\int_{x_b}^x \gamma_0(y)dy\right) \left(\frac{1}{d}\right)^{k-1} (a\zeta + b)^{j-1} (c\zeta + d)^{k-j}.\end{aligned}$$

So apart from the factor  $\exp\left(\int_{x_b}^x \gamma_0(y)dy\right)$  the components  $\tilde{w}_j(x)$  are linear combinations of the powers  $\zeta^\ell$ ,  $\ell = 0, 1, \dots, k-1$ . We conclude that the requirement that  $\{w_1, w_2, \dots, w_k\}$  and  $\{\tilde{w}_1, \tilde{w}_2, \dots, \tilde{w}_k\}$  are equivalent systems of output functionals is indeed satisfied.

A transformation  $y = \phi(x)$  leaves the form of  $g$  invariant. Because of the extra factor (recall (5.4)) we have to adapt  $\gamma_0$  by a factor, too. The net effect is that the integral of  $\gamma_0$ , and hence  $w$ , transforms in the standard manner.

By choosing  $y = \zeta(x)$  as a new variable we eliminate the factor  $v_0(x)$  and  $g$  becomes a quadratic polynomial in the (transformed) i-state variable. If we subsequently take  $t = \log \zeta$  as the new i-state variable, we find that the growth rate of  $t$  has the form

$$v_1 e^{-t} + v_2 + v_3 e^t.$$

If, alternatively, we take  $t = 2 \arctan \zeta$  as the new  $i$ -state variable, we can use the identities

$$\begin{aligned}\cos t &= \frac{1 - \zeta^2}{1 + \zeta^2}, \\ \sin t &= \frac{2\zeta}{1 + \zeta^2}, \\ \frac{d}{d\zeta} \arctan \zeta &= \frac{1}{1 + \zeta^2}\end{aligned}$$

to find that the growth rate of  $t$  now has the form

$$v_1(1 + \cos t) + v_2 \sin t + v_3(1 - \cos t).$$

We have presented the two transformations above for two reasons:

- they illustrate that appearances are deceptive (a growth rate may at first sight seem to fail to fit into the catalogue, while in fact it does),
- they play a role in the proofs.

We are now ready to state our main results.

**Theorem 5.2** *If  $g$ ,  $\mu$  and  $w$  are of the form specified in either  $F_2$  or  $F_3$ , then (4.33) holds for the indicated matrix  $H(E)$  and  $w_1, w_2, \dots, w_k$  are linearly independent functions of  $x \in \Omega$ .*

**Theorem 5.3** *Assume that  $k \geq 2$  and that*

- (i)  *$g$ ,  $\mu$ ,  $w$  and  $H$  are such that (4.33) holds,*
- (ii)  *$w_1, w_2, \dots, w_k$  are linearly independent functions of  $x \in \Omega$ .*

*Then necessarily  $g$ ,  $\mu$  and  $w$  can be brought into the form specified in either  $F_2$  or  $F_3$  by a transformation of the  $i$ -state variable and a change of basis in  $\mathbb{R}^k$ .*

The proofs of these two theorems are given in the next section.

## 6 Proofs

### 6.1 Some spadework

As a prelude we recall some standard theory concerning linear ODE-systems (cf. Hirsch and Smale 1974). Let  $A$  be a real  $k \times k$  matrix. Then  $A$  defines a linear map from  $\mathbb{C}^k$  to  $\mathbb{C}^k$  that leaves  $\mathbb{R}^k$ , considered as the linear subspace of  $\mathbb{C}^k$  consisting of vectors with zero imaginary part, invariant. Let  $\phi_{j,1}$ ,  $j = 1, 2, \dots, r$  be a maximal set of linearly independent eigenvectors of  $A$  and let  $\lambda_1, \lambda_2, \dots, \lambda_r$  be the corresponding eigenvalues. Thus the geometric multiplicity of an eigenvalue determines how many times it is listed. We choose  $\phi_{j,1} \in \mathbb{R}^k$  if  $\lambda_j$  is real and  $\phi_{j_2,1} = \overline{\phi_{j_1,1}}$  if  $(\lambda_{j_1}, \lambda_{j_2})$  is a pair of non-real complex conjugate eigenvalues. Let  $k_1, k_2, \dots, k_r$  be the corresponding

multiplicities, that is,  $k_j$  is the length of the Jordan block generated by  $\phi_{j,1}$ . For  $j = 1, 2, \dots, r$  and  $\ell = 1, 2, \dots, k_j$ , let  $\phi_{j,\ell}$  be such that, with the convention  $\phi_{j,0} = 0$ ,

$$\Lambda \phi_{j,\ell} = \lambda_j \phi_{j,\ell} + \phi_{j,\ell-1} \quad (6.1)$$

and, moreover,  $\phi_{j,\ell} \in \mathbb{R}^k$  if  $\lambda_j$  is real and  $\phi_{j_2,\ell} = \overline{\phi_{j_1,\ell}}$  if  $(\lambda_{j_1}, \lambda_{j_2})$  is a pair of non-real complex conjugate eigenvalues. Then for the complex variable  $z$  we have

$$e^{z\Lambda} \phi_{j,\ell} = e^{\lambda_j z} \left( \phi_{j,\ell} + \sum_{m=1}^{\ell-1} \frac{z^m}{m!} \phi_{j,\ell-m} \right). \quad (6.2)$$

The identity

$$\sum_{j=1}^r k_j = k \quad (6.3)$$

holds and the set  $\{\phi_{j,\ell}\}$  is a basis for  $\mathbb{C}^k$ . So for any  $\psi(0) \in \mathbb{C}^k$  there exist coefficients  $c_{j,\ell} \in \mathbb{C}$  such that

$$\psi(0) = \sum_{j,\ell} c_{j,\ell} \phi_{j,\ell} \quad (6.4)$$

and hence

$$e^{z\Lambda} \psi(0) = \sum_{j,\ell} c_{j,\ell} e^{\lambda_j z} \left( \phi_{j,\ell} + \sum_{m=1}^{\ell-1} \frac{z^m}{m!} \phi_{j,\ell-m} \right). \quad (6.5)$$

When  $\psi(0) \in \mathbb{R}^k$  we have  $c_{j,\ell} \in \mathbb{R}$  if  $\lambda_j$  is real and  $c_{j_2,\ell} = \overline{c_{j_1,\ell}}$  if  $(\lambda_{j_1}, \lambda_{j_2})$  is a pair of non-real complex conjugate eigenvalues. If we restrict the variable  $z$  to the real axis we denote it by  $\xi$ . We have  $\exp(\xi \Lambda) \psi(0) \in \mathbb{R}^k$  if  $\psi(0) \in \mathbb{R}^k$  even though the representation (6.5) possibly involves complex quantities. We want to determine conditions on  $\Lambda$  and  $\psi(0)$  that guarantee that the components of  $\exp(\xi \Lambda) \psi(0)$ , regarded as real valued functions of the real variable  $\xi$  are linearly independent. To derive such conditions we first investigate the complex setting.

A preliminary yet key observation is that the functions

$$z \mapsto z^{n_1} e^{\lambda_1 z} \quad \text{and} \quad z \mapsto z^{n_2} e^{\lambda_2 z}$$

are linearly dependent as complex functions if and only if  $\lambda_1 = \lambda_2$  and  $n_1 = n_2$ . This is obvious, because the Wronskian determinant of the two functions vanishes identically if and only if the latter condition is satisfied. By the same argument one obtains the analogous results for a finite collection of functions of the form  $z \mapsto z^n \exp(\lambda z)$ . Recall that for *analytic* functions the identical vanishing of the Wronskian is necessary and sufficient for linear dependence (Bôcher 1900).

**Lemma 6.1** *The components of*

$$\psi(z) = e^{z\Lambda} \psi(0) \quad (6.6)$$

*are linearly independent functions of the complex variable  $z$  if and only if the following two conditions are met*

- (i) *the eigenvalues of  $\Lambda$  have geometric multiplicity one,*
- (ii)  *$c_{j,k_j} \neq 0$ ,  $j = 1, 2, \dots, r$  if  $\psi(0)$  is represented by (6.4).*

**Proof** As (6.5) clearly shows, each component of  $\psi(z)$  is a linear combination of building blocks of the form  $z^m \exp(\lambda_j z)$ . According to (6.3) there are exactly  $k$  building blocks. To make  $k$  linearly independent linear combinations we need  $k$  linearly independent building blocks. It follows that the conditions (i) and (ii) are necessary.

By definition the components of  $\psi(z)$  are linearly independent if

$$d \in \mathbb{C}^k, \quad d \cdot \psi(z) = 0 \quad \text{for all } z \in \mathbb{C} \quad \Rightarrow \quad d = 0. \quad (6.7)$$

Since  $\{\phi_{j,\ell}\}$  is a basis for  $\mathbb{C}^k$ , we have  $d = 0$  if and only if  $d \cdot \phi_{j,\ell} = 0$  for  $j = 1, 2, \dots, r$ ,  $\ell = 1, 2, \dots, k_j$ . So we want to show that when (i) and (ii) hold, then

$$d \cdot \psi(z) = 0 \quad \text{for all } z \in \mathbb{C} \quad \Rightarrow \quad d \cdot \phi_{j,\ell} = 0 \quad \text{for all } j, \ell. \quad (6.8)$$

To do so, we deduce from (6.5) that

$$d \cdot \psi(z) = \sum_{j,\ell} c_{j,\ell} e^{\lambda_j z} \left( d \cdot \phi_{j,\ell} + \sum_{m=1}^{\ell-1} \frac{z^m}{m!} d \cdot \phi_{j,\ell-m} \right). \quad (6.9)$$

Because of condition (i) the right hand side of (6.9) can be identically zero only if for  $j = 1, 2, \dots, r$

$$\sum_{\ell=1}^{k_j} c_{j,\ell} \left( d \cdot \phi_{j,\ell} + \sum_{m=1}^{\ell-1} \frac{z^m}{m!} d \cdot \phi_{j,\ell-m} \right) = 0 \quad \text{for all } z \in \mathbb{C}, \quad (6.10)$$

or, equivalently,

$$\sum_{m=0}^{k_j-1} \left( \sum_{\ell=m+1}^{k_j} c_{j,\ell} d \cdot \phi_{j,\ell-m} \right) \frac{z^m}{m!} = 0 \quad \text{for all } z \in \mathbb{C}. \quad (6.11)$$

Condition (6.11) in turn leads to

$$\sum_{\ell=m+1}^{k_j} c_{j,\ell} d \cdot \phi_{j,\ell-m} = 0, \quad m = 0, 1, \dots, k_j - 1. \quad (6.12)$$

Now assume that condition (ii) holds. Taking  $m = k_j - 1$  we find that necessarily  $d \cdot \phi_{j,1} = 0$ . Next, taking  $m = k_j - 2$  and using  $d \cdot \phi_{j,1} = 0$ , we find that  $d \cdot \phi_{j,2} = 0$ . Continuing in this way we prove (6.8) and consequently (6.7) holds. This completes the sufficiency part of the proof.  $\square$

**Lemma 6.2** *Let  $f_j$ ,  $j = 1, 2, \dots, k$ , be analytic functions defined in an open subset  $A$  of  $\mathbb{C}$  that contains an interval of  $\mathbb{R}$  and assume that  $f_j(A \cap \mathbb{R}) \subset \mathbb{R}$ . Then the functions  $f_1, f_2, \dots, f_k$  are linearly independent as functions of a complex variable if and only if their restrictions to  $A \cap \mathbb{R}$  are linearly independent as functions of a real variable.*

**Proof** This follows immediately from the fact that  $k$  analytic functions are linearly dependent if and only if their Wronskian determinant vanishes identically (Böcher 1900).  $\square$

**Corollary 6.3** *Let  $\Lambda$  be a real  $k \times k$  matrix and let  $\psi(0) \in \mathbb{R}^k$ . The components of*

$$\psi(\xi) = e^{\xi \Lambda} \psi(0) \quad (6.13)$$

*are linearly independent functions of the real variable  $\xi$  if and only if the following conditions are met*

- (i) *the eigenvalues of  $\Lambda$  have geometric multiplicity one,*
- (ii) *if  $\psi(0)$  is expressed as a linear combination of generalised eigenvectors of  $\Lambda$ , then the coefficients of generalized eigenvectors of highest rank are non-zero (cf. (6.4) and condition (ii) of Lemma 6.1).*

Let  $\Lambda$  and  $\psi(0)$  satisfy the conditions (i) and (ii) of Corollary 6.3. If we order  $\{\xi^m \exp(\lambda_j \xi)\}$  (say lexicographically) as  $\{h_\ell(\xi)\}_{\ell=1}^k$ , then there exist coefficients  $d_{j\ell}$  such that

$$\psi_j(\xi) = \sum_{\ell=1}^k d_{j\ell} h_\ell(\xi). \quad (6.14)$$

Let  $c \in \mathbb{R}^k$ . Then

$$c \cdot \psi(\xi) = \sum_{j=1}^k c_j \psi_j(\xi) = \sum_{\ell=1}^k \left( \sum_{j=1}^k c_j d_{j\ell} \right) h_\ell(\xi). \quad (6.15)$$

By the linear independence of  $\{h_\ell\}_{\ell=1}^k$  we have that  $c \cdot \psi = 0$  if and only if  $\sum_{j=1}^k c_j d_{j\ell} = 0$  for  $\ell = 1, 2, \dots, k$ . By Corollary 6.3  $c \cdot \psi = 0$  implies  $c = 0$ . This translates into

$$cD = 0 \quad \Rightarrow \quad c = 0, \quad (6.16)$$

where  $D$  is the  $k \times k$  matrix with elements  $d_{j\ell}$ . We conclude that  $D$  is invertible. So when conditions (i) and (ii) of Corollary 6.3 hold, there exists an invertible transformation  $D$  that relates the vector  $\psi(\xi) = \exp(\xi \Lambda) \psi(0)$  to the vector  $h(\xi)$  with components  $h_\ell(\xi)$ .

Note that  $h_{\ell_1}$  is complex valued if  $\operatorname{Im} \lambda_{j_1} \neq 0$ , but that in that case there exists  $\ell_2$  such that  $h_{\ell_2}(\xi) = \bar{h}_{\ell_1}(\xi)$  and, moreover,  $d_{j\ell_2} = \bar{d}_{j\ell_1}$ .

## 6.2 Proof of Theorem 5.2

### 6.2.1 The model family $F_2$

The proof that (4.33) holds for  $F_2$  amounts to a straightforward verification. The linear independence of  $\{w_1, w_2, \dots, w_k\}$  is guaranteed by Corollary 6.3.  $\square$

### 6.2.2 The model family $F_3$

Let  $g, \mu, w, \zeta$  be as specified in  $F_3$ . Since

$$\zeta(x) = \int_{x_b}^x \frac{dy}{v_0(y)},$$

$\zeta(x)$  is not constant and as a consequence the components

$$w_j(x) = \exp\left(\int_{x_b}^x \gamma_0(y) dy\right) \zeta(x)^{j-1} \quad (6.17)$$

of  $w$  are linearly independent functions of  $x$ .

To verify (4.33), which under  $F_3$  takes the form

$$g(x, E) \frac{dw}{dx}(x) = H_0(E)w(x) + (\mu(x, E) - \mu_0(E))w(x), \quad (6.18)$$

note that

$$\frac{d}{dx} (\zeta(x))^{j-1} = (j-1) (\zeta(x))^{j-2} \frac{1}{v_0(x)} \quad (6.19)$$

and hence

$$\begin{aligned} g(x, E) \frac{d}{dx} (\zeta(x))^{j-1} &= (j-1) \left( v_1(E) (\zeta(x))^{j-2} + v_2(E) (\zeta(x))^{j-1} \right. \\ &\quad \left. + v_3(E) (\zeta(x))^j \right). \end{aligned} \quad (6.20)$$

Differentiating (6.17) one obtains

$$\frac{d}{dx} w_j(x) = \gamma_0(x) w_j(x) + e^{\int_{x_b}^x \gamma_0(y) dy} \frac{d}{dx} (\zeta(x))^{j-1} \quad (6.21)$$



and multiplying (6.21) by  $g(x, E)$ , taking (6.17), (6.19) and the form of  $\mu(x, E)$  specified in  $F_3$  into account, one obtains

$$\begin{aligned}
 g(x, E) \frac{d}{dx} w_j(x) &= \gamma_0(x) g(x, E) w_j(x) \\
 &\quad + e^{\int_{x_b}^x \gamma_0(y) dy} (j-1) \left( v_1(E) \zeta(x)^{j-2} \right. \\
 &\quad \left. + v_2(E) \zeta(x)^{j-1} + v_3(E) \zeta(x)^j \right) \\
 &= (\mu(x, E) - \mu_0(E)) w_j(x) \\
 &\quad + e^{\int_{x_b}^x \gamma_0(y) dy} \left( (j-1) v_1(E) \zeta(x)^{j-2} + (j-1) v_2(E) \zeta(x)^{j-1} \right. \\
 &\quad \left. - (k-j) v_3(E) \zeta(x)^j \right) \tag{6.22}
 \end{aligned}$$

The  $j$ th component of the right hand side of (6.18) and the right hand side of (6.22) contain the common term  $(\mu(x, E) - \mu_0(E)) w_j(x)$ . Thus to verify that (6.18) holds, we only have to check that the  $j$ th component of the vector obtained by applying the matrix  $H_0(E)$  as specified in  $F_3$  to the vector  $(1 \ \zeta(x) \ \zeta(x)^2 \ \dots \ \zeta(x)^{k-1})^T$  is

$$(j-1) v_1(E) \zeta(x)^{j-2} + (j-1) v_2(E) \zeta(x)^{j-1} - (k-j) v_3(E) \zeta(x)^j$$

and this is obviously the case.  $\square$

### 6.3 Proof of Theorem 5.3

For  $k = 2$  and for given  $H$  and  $w$  we may consider (4.33) as two linear equations in the two unknowns  $g$  and  $\mu$ . The solution has the form given in  $F_3$  with  $\zeta$ ,  $v_i$ ,  $\gamma_0$  and  $\mu_0$  expressed in terms of  $w$  and  $H$ . This computation provides a strong indication that Theorem 5.3 is correct, but it does not yield a proof.

Our strategy is to focus first on the fixed value  $E_0$  of  $E$  such that  $g(x, E_0) > 0$  for all  $x \in \Omega$ . The existence of such an  $E_0$  is guaranteed by Assumption 5.1. This enables us to show that  $w$  is a transformed version of a matrix exponential and hence consists of building blocks of the form  $\xi^m \exp(\lambda \xi)$ . After that we consider general  $E$  and investigate the consequences for (4.33).

Define

$$V(x) := \exp \left( - \int_{x_b}^x \frac{\mu(y, E_0)}{g(y, E_0)} dy \right) w(x). \tag{6.23}$$

Then (4.33) is equivalent to

$$g(x, E) V'(x) = (H(E) + \tilde{\mu}(x, E)) V(x) \tag{6.24}$$

with

$$\tilde{\mu}(x, E) := \mu(x, E) - \frac{g(x, E)}{g(x, E_0)} \mu(x, E_0) \quad (6.25)$$

and hence

$$\tilde{\mu}(x, E_0) = 0 \quad \text{for all } x \in \Omega. \quad (6.26)$$

Next define

$$\tau(x) := \int_{x_b}^x \frac{dy}{g(y, E_0)} \quad (6.27)$$

and

$$\Phi(t) := V\left(\tau^{-1}(t)\right). \quad (6.28)$$

Then (6.24) is equivalent to

$$\tilde{g}(t, E)\Phi'(t) = \left(H(E) + \tilde{\mu}\left(\tau^{-1}(t), E\right)\right)\Phi(t) \quad (6.29)$$

with

$$\tilde{g}(t, E) = \frac{g\left(\tau^{-1}(t), E\right)}{g\left(\tau^{-1}(t), E_0\right)} \quad (6.30)$$

and hence

$$\tilde{g}(t, E_0) = 1 \quad \text{for all } t \in \tau(\Omega). \quad (6.31)$$

It follows that

$$\Phi(t) = e^{tH(E_0)}\Phi(0) \quad (6.32)$$

and hence necessarily  $H(E_0)$  and  $\Phi(0)$  satisfy the conditions (i) and (ii) of Corollary 6.3 (as otherwise  $w$  would not consist of  $k$  linearly independent functions of  $x$ ).

Let  $\lambda_1, \lambda_2, \dots, \lambda_r$  denote the eigenvalues of  $H(E_0)$  with multiplicities  $k_1, k_2, \dots, k_r$  such that  $\sum_{p=1}^r k_p = k$ . If we order  $\{t^m \exp(\lambda_p t)\}$  lexicographically as  $\{h_\ell(t)\}_{\ell=1}^k$ , the analogue of (6.14) reads

$$\Phi(t) = DX(t). \quad (6.33)$$

Applying the matrix  $D^{-1}$  to (6.29) we find

$$\tilde{g}(t, E)X'(t) - \tilde{\mu}\left(\tau^{-1}(t), E\right)X(t) = D^{-1}H(E)DX(t). \quad (6.34)$$

Every component of the right hand side of (6.34) is a linear combination, with  $E$ -dependent coefficients, of the building blocks

$$h_\ell(t) = t^m e^{\lambda_p t}. \quad (6.35)$$

The  $\ell$ th component of the left hand side of (6.34) reads

$$\tilde{g}(t, E) \left( mt^{m-1} + \lambda_p t^m \right) e^{\lambda_p t} - \tilde{\mu} \left( \tau^{-1}(t), E \right) t^m e^{\lambda_p t}.$$

Hence (6.34) implies that for  $p = 1, 2, \dots, r$ ,  $m = 0, 1, \dots, k_p - 1$  one has

$$\left( mt^{m-1} + \lambda_p t^m \right) \tilde{g}(t, E) - t^m \tilde{\mu} \left( \tau^{-1}(t), E \right) = \sum_{j=1}^r \sum_{\ell=0}^{k_j-1} c_{j\ell} t^\ell e^{(\lambda_j - \lambda_p)t}, \quad (6.36)$$

where the coefficients  $c_{j\ell}$  depend on  $E$ ,  $m$  and  $p$ .

Apart from  $\tilde{g}$ ,  $\tilde{\mu}$  and powers of  $t$ , the identity (6.36) involves exponential functions with exponents from the set

$$U_p := \{ \lambda_j - \lambda_p : j = 1, 2, \dots, r \}. \quad (6.37)$$

Since (6.36) should hold for  $p = 1, 2, \dots, r$ , we are particularly interested in a characterisation of

$$\bigcap_{p=1}^r U_p.$$

As already observed above, we know that necessarily

$$\lambda_j \neq \lambda_\ell \quad \text{if } j \neq \ell. \quad (6.38)$$

**Lemma 6.4** *Let  $r \geq 2$ . Then*

$$\bigcap_{p=1}^r U_p = \{0\}.$$

**Proof** Let  $\alpha \in \bigcap_{p=1}^r U_p$  and assume that  $\alpha \neq 0$ . Since  $\alpha \in U_j$  there exists an  $m = m(j)$  such that

$$\lambda_m = \lambda_j + \alpha. \quad (6.39)$$

Because of (6.38), there is at most one  $m$  for which (6.39) holds. We write  $m = s(j)$ , with  $s$  standing for “successor”, and thus define a map  $s : \{1, 2, \dots, r\} \rightarrow$

$\{1, 2, \dots, r\}$ . The sequence  $1, s(1), s^2(1), \dots$  takes values in a finite set so necessarily some value is taken at least twice. This implies that

$$s^\ell(j) = j \quad (6.40)$$

for some  $j$  and  $\ell \geq 1$ . But (6.40) means that

$$\lambda_j = \lambda_j + \ell\alpha \quad (6.41)$$

so after all,  $\alpha = 0$ .  $\square$

Suppose that, with appropriate numbering of the eigenvalues of  $H(E_0)$ , we have

$$\lambda_j = \lambda_1 + (j - 1)\alpha. \quad (6.42)$$

Then

$$U_p = \{(j - p)\alpha : j = 1, 2, \dots, r\} \quad (6.43)$$

and hence

$$\bigcap_{p=2}^r U_p = \{-\alpha, 0\}. \quad (6.44)$$

We now show that the existence of a non-zero element  $-\alpha$  in  $\bigcap_{p=2}^r U_p$  implies that (6.42) holds, or, in other words, that (6.42) and (6.44) are equivalent.

**Lemma 6.5** *Let  $r \geq 2$ . Either  $\bigcap_{p=2}^r U_p = \{0\}$  or, possibly after renumbering the  $\lambda$ 's, (6.42) holds.*

**Proof** Let  $-\alpha \in \bigcap_{p=2}^r U_p$  and assume that  $\alpha \neq 0$ . As in the proof of Lemma 6.4 we define the successor map  $s$ . But now  $s$  is defined on  $\{2, 3, \dots, r\}$  while taking values in  $\{1, 2, \dots, r\}$ . If the sequence  $j, s(j), s^2(j), \dots$  takes values only in  $\{2, 3, \dots, r\}$ , it necessarily becomes periodic and again we conclude that actually  $\alpha = 0$ . Hence  $\alpha \neq 0$  requires that the sequence hits 1 after finitely many steps, implying that

$$\lambda_1 = \lambda_j - m(j)\alpha$$

which with appropriate numbering of the  $\lambda$ 's amounts to (6.42).  $\square$

In our analysis of (6.36) we shall also subtract the identity with  $p = p_1$  from the identity with  $p = p_2$ . This motivates us to consider the intersection of all unions  $U_{p_1} \cup U_{p_2}$  with  $p_1 \neq p_2$ .

**Lemma 6.6** *If*

$$\alpha \in \bigcap_{p_1 \neq p_2} (U_{p_1} \cup U_{p_2}) \setminus \{0\} \quad (6.45)$$

then, with appropriate numbering of the  $\lambda$ 's, (6.42) holds and

$$\bigcap_{p_1 \neq p_2} (U_{p_1} \cup U_{p_2}) = \{-\alpha, 0, \alpha\}. \quad (6.46)$$

**Proof** Assume (6.45). We claim that there can be at most one index  $p$  such that  $\alpha \notin U_p$ , that is,

$$\alpha \neq \lambda_j - \lambda_p \quad \text{for } j = 1, 2, \dots, r. \quad (6.47)$$

Indeed, if (6.47) would hold for  $p = q_1$  as well as for  $p = q_2$ , with  $q_1 \neq q_2$ , then

$$\alpha \notin U_{q_1} \cup U_{q_2}$$

and hence certainly

$$\alpha \notin \bigcap_{p_1 \neq p_2} (U_{p_1} \cup U_{p_2}).$$

Lemma 6.4 implies that at least one index  $p$  must exist such that  $\alpha \notin U_p$ . So there exists a unique index  $p$  such that  $\alpha \notin U_p$ . Let us call it the exceptional index.

Renumber the  $\lambda$ 's such that the exceptional index is 1. Then  $\alpha \in \bigcap_{p=2}^r U_p$  and we can apply Lemma 6.5 to deduce that (6.42) holds. It follows that elements of

$$\bigcap_{p_1 \neq p_2} (U_{p_1} \cup U_{p_2})$$

are of the form  $m\alpha$  with  $1 - r \leq m \leq r - 1$ . The number 0 belongs to  $U_p$  for all  $p$ , the number  $\alpha$  for all  $p \neq r$ , and the number  $-\alpha$  for all  $p \neq 1$ . So the right hand side of (6.46) is a subset of the left hand side. If  $m \geq 2$ , the number  $m\alpha$  does not belong to  $U_p$  for  $p = r - 1, r$  and if  $m \leq -2$  it does not belong to  $U_p$  for  $p = 1, 2$ . We conclude that (6.46) holds.  $\square$

**Lemma 6.7** Let  $r \geq 2$  and assume that

$$\bigcap_{p_1 \neq p_2} (U_{p_1} \cup U_{p_2}) = \{0\}. \quad (6.48)$$

Then both  $t \mapsto \tilde{g}(t, E)$  and  $t \mapsto \tilde{\mu}(\tau^{-1}(t), E)$  are constant functions.

**Proof** Consider (6.36) with  $m = 0$  for  $p = p_1$  and for  $p = p_2$ . By subtraction we find

$$(\lambda_{p_2} - \lambda_{p_1}) \tilde{g}(t, E) = \sum_{j=1}^r \sum_{\ell=0}^{k_j-1} \left( c_{j\ell}(p_2) e^{(\lambda_j - \lambda_{p_2})t} - c_{j\ell}(p_1) e^{(\lambda_j - \lambda_{p_1})t} \right) t^\ell. \quad (6.49)$$

The identity (6.49) holds for all  $p_1$  and  $p_2$  with  $p_1 \neq p_2$  (and of course trivially for  $p_1 = p_2$ ). It therefore follows from (6.48) that all exponentials with non-zero exponent in (6.49) must have coefficient zero, that is, necessarily

$$c_{j\ell}(p_m) = 0 \quad \text{if } j \neq p_m.$$

It follows that the map  $t \mapsto \tilde{g}(t, E)$  is a polynomial of degree  $d_1 \leq \max\{k_p\} - 1$ .

Returning to (6.36) with  $m = 0$ , we express  $\tilde{\mu}$  in terms of  $\lambda_p \tilde{g}$  and the sum at the right hand side. From the arbitrariness of  $p$  and Lemma 6.4 we deduce that the map  $t \mapsto \tilde{\mu}(\tau^{-1}(t), E)$ , too, is a polynomial and that its degree  $d_2$  does not exceed  $\max\{k_p\} - 1$ . Hence the map

$$t \mapsto \left( \lambda_p \tilde{g}(t, E) - \tilde{\mu}(\tau^{-1}(t), E) \right) t^m + m t^{m-1} \tilde{g}(t, E)$$

is a polynomial of degree  $d_3 = \max\{d_1, d_2\} + m$  (unless  $\lambda_p \tilde{g} - \tilde{\mu} = 0$ , but this can be the case for at most one value of  $p$ ).

The right hand side of (6.36) is a polynomial in  $t$  if and only if  $c_{j\ell}(p) = 0$  for  $j \neq p$  and, if it is, the degree is at most  $k_p - 1$ . We conclude that

$$\max\{d_1, d_2\} + m \leq k_p - 1$$

and consequently, by taking  $m = k_p - 1$ , that

$$\max\{d_1, d_2\} = 0.$$

□

**Corollary 6.8** *Let  $r \geq 2$  and assume that (6.48) holds. Then  $g$ ,  $\mu$ ,  $w$  and  $H$  are of the form specified in  $F_2$  with*

$$\begin{aligned} v_0(x) &= g(x, E_0), \\ v_1(E) &= \tilde{g}(\tau(x), E), \\ \gamma_0(x) &= \frac{\mu(x, E_0)}{g(x, E_0)}, \\ \mu_0(E) &= \tilde{\mu}(x, E), \\ \zeta(x) = \tau(x) &= \int_{x_b}^x \frac{dy}{g(y, E_0)}, \\ \Lambda &= H(E_0). \end{aligned}$$

**Proof** From (6.30) we know that

$$g(x, E) = g(x, E_0) \tilde{g}(\tau(x), E) \tag{6.50}$$

and by Lemma 6.7 the second factor on the right hand side of (6.50) is constant as a function of its first argument, so with the given definitions of  $v_0$  and  $v_1$ , the function  $g$  is indeed of the form specified in  $F_2$ .

Likewise it follows from (6.25) that

$$\mu(x, E) = \frac{\mu(x, E_0)}{g(x, E_0)} g(x, E) + \tilde{\mu}(x, E). \quad (6.51)$$

The last term on the right hand side of (6.51) is, by Lemma 6.7, constant as a function of its first argument and hence  $\mu$ , too, is of the form specified in  $F_2$  with  $\gamma_0$  and  $\mu_0$  as listed.

From (6.23), (6.28) and (6.32) we deduce that

$$w(x) = e^{\int_{x_b}^x \gamma_0(y) dy} \Phi((\tau(x))) = e^{\int_{x_b}^x \gamma_0(y) dy} e^{\tau(x)H(E_0)} \Phi(0), \quad (6.52)$$

where  $\tau(x)$  is defined by (6.27). Now recall (6.31) and translate it into  $v_1(E_0) = 1$ . It follows that

$$\tau(x) = \int_{x_b}^x \frac{dy}{g(y, E_0)} = \int_{x_b}^x \frac{dy}{v_0(y)} \quad (6.53)$$

Substituting (6.53) into (6.52) and taking into account that  $\Lambda = H(E_0)$  because  $v_1(E_0) = 1$  and  $\mu_0(E_0) = 0$  by (6.26), we see that  $w$  is of the form specified in  $F_2$ .  $\square$

**Lemma 6.9** *Let  $r = 1$ ,  $k \geq 2$ . Then necessarily*

$$\tilde{g}(t, E) = v_1(E) + v_2(E)t + v_3(E)t^2, \quad (6.54)$$

$$\tilde{\mu}(\tau^{-1}(t), E) = \lambda \tilde{g}(t, E) + \mu_0(E) + (k-1)v_3(E)t. \quad (6.55)$$

**Proof** Because we assume that  $r = 1$ , we can write (6.36) as

$$\left( \lambda \tilde{g}(t, E) - \tilde{\mu}(\tau^{-1}(t), E) \right) t^m + m t^{m-1} \tilde{g}(t, E) = \sum_{\ell=0}^{k-1} c_\ell t^\ell, \quad (6.56)$$

with  $c_\ell$  depending on  $E$  and  $m$ . We make the dependence on  $m$  explicit by writing  $c_\ell(m)$ .

If we take  $m = 0$  and multiply (6.56) by  $t$  we obtain

$$\left( \lambda \tilde{g}(t, E) - \tilde{\mu}(\tau^{-1}(t), E) \right) t = \sum_{\ell=0}^{k-1} c_\ell(0) t^{\ell+1} \quad (6.57)$$

while for  $m = 1$ , (6.56) takes the form

$$\left( \lambda \tilde{g}(t, E) - \tilde{\mu}(\tau^{-1}(t), E) \right) t + \tilde{g}(t, E) = \sum_{\ell=0}^{k-1} c_\ell(1) t^\ell. \quad (6.58)$$

Subtracting (6.57) from (6.58) we find that  $\tilde{g}(t, E)$  is a polynomial in  $t$ .

Next we observe that (6.56) with  $m = k - 2$  and multiplied by  $t$  becomes

$$\left(\lambda \tilde{g}(t, E) - \tilde{\mu}(\tau^{-1}(t), E)\right) t^{k-1} + (k-2)t^{k-2}\tilde{g}(t, E) = \sum_{\ell=0}^{k-1} c_{\ell} t^{\ell+1} \quad (6.59)$$

and that for  $m = k - 1$  the identity (6.56) reads

$$\left(\lambda \tilde{g}(t, E) - \tilde{\mu}(\tau^{-1}(t), E)\right) t^{k-1} + (k-1)t^{k-2}\tilde{g}(t, E) = \sum_{\ell=0}^{k-1} c_{\ell} t^{\ell}. \quad (6.60)$$

If we subtract (6.60) from (6.59), we find that  $t^{k-2}\tilde{g}(t, E)$  is a polynomial in  $t$  of degree at most  $k$ . Since  $\tilde{g}(t, E)$  itself is a polynomial in  $t$ , we conclude that its degree is at most 2, that is, (6.54) must hold.

It follows from (6.54) and (6.60) that

$$\left(\lambda \tilde{g}(t, E) - \tilde{\mu}(\tau^{-1}(t), E)\right) t^{k-1}$$

is a polynomial of degree at most  $k$  and hence

$$\lambda \tilde{g}(t, E) - \tilde{\mu}(\tau^{-1}(t), E) = -\mu_0(E) - \mu_1(E)t. \quad (6.61)$$

Substituting (6.61) into (6.60) shows that necessarily

$$\mu_1(E) = (k-1)v_3(E).$$

This yields (6.55).  $\square$

We say that a square matrix is in *subdiagonal Jordan form* if the entries on the main diagonal equal  $\lambda$  and the entries on the subdiagonal (that is, the entries immediately under the diagonal) equal 1 and all other entries equal 0.

**Corollary 6.10** *Let  $r = 1$ ,  $k \geq 2$ . Assume that  $H(E_0)$  is in subdiagonal Jordan form and let  $\Phi(0) = (1 \ 0 \ \cdots \ 0)^T$ . Then  $g$ ,  $\mu$ ,  $w$  and  $H$  are as specified in  $F_3$ .*

**Proof** Because  $H(E_0)$  is in subdiagonal Jordan form and  $\Phi(0) = (1 \ 0 \ \cdots \ 0)^T$ , we have

$$\Phi(t) = e^{\lambda t} \begin{pmatrix} 1 \\ t \\ \vdots \\ t^{k-1} \end{pmatrix}. \quad (6.62)$$

Combining (6.30) and (6.54) we find that  $g$  has the form specified in  $F_3$  with

$$v_0(x) = g(x, E_0)$$



and hence

$$\zeta(x) = \tau(x),$$

where  $\tau(x)$  is defined by (6.27). Combining (6.25) and (6.55) we find that  $\mu$  is of the form specified in  $F_3$  with

$$\gamma_0(x) = \frac{\mu(x, E_0) + \lambda}{g(x, E_0)}.$$

Combining (6.23), (6.28) and (6.62) and taking (6.27) into account, we see that  $w$  is of the form specified in  $F_3$ .

Finally we deduce the form of  $H(E)$  by substituting what we know about  $g$ ,  $\mu$  and  $w$  into (4.33).  $\square$

**Lemma 6.11** *Let  $r \geq 2$  and assume that*

$$\alpha \in \bigcap_{p_1 \neq p_2} (U_{p_1} \cup U_{p_2}) \setminus \{0\}.$$

*Then necessarily  $k_j = 1$  for  $j = 1, 2, \dots, r$  and hence  $r = k$  and, moreover, there exist  $u_j(E)$ ,  $j = 1, 2, 3, 4$  such that*

$$\tilde{g}(t, E) = u_1(E)e^{-\alpha t} + u_2(E) + u_3(E)e^{\alpha t}, \quad (6.63)$$

$$\tilde{\mu}(\tau^{-1}(t), E) = \lambda_1 u_1(E)e^{-\alpha t} + u_4(E) + \lambda_k u_3(E)e^{\alpha t}, \quad (6.64)$$

*where the  $\lambda$ 's are numbered such that (6.42) holds. One has*

$$u_1(E_0) = 0, \quad u_2(E_0) = 1, \quad u_3(E_0) = 0, \quad u_4(E_0) = 0. \quad (6.65)$$

**Proof** Recall that in Lemma 6.6 we established that, with appropriate numbering, (6.42) holds and that

$$\bigcap_{p_1 \neq p_2} (U_{p_1} \cup U_{p_2}) = \{-\alpha, 0, \alpha\}.$$

From (6.49) we now deduce that

$$\tilde{g}(t, E) = P_-(t, E)e^{-\alpha t} + P_0(t, E) + P_+(t, E)e^{\alpha t},$$

where  $P_-$ ,  $P_0$  and  $P_+$  are all polynomials in  $t$  of degree at most  $\max\{k_p\} - 1$ . Next (6.36) with  $m = 0$  and Lemma 6.4 imply that

$$\tilde{\mu}(\tau^{-1}(t), E) = Q_-(t, E)e^{-\alpha t} + Q_0(t, E) + Q_+(t, E)e^{\alpha t},$$

where, likewise,  $Q_-$ ,  $Q_0$  and  $Q_+$  are polynomials in  $t$  of degree at most  $\max\{k_p\} - 1$ .

We rewrite (6.36) with  $p = 1$  and  $m \leq k_1 - 1$  as

$$\begin{aligned} & t^m \left( (\lambda_1 P_-(t, E) - Q_-(t, E))e^{-\alpha t} + \lambda_1 P_0(t, E) - Q_0(t, E) \right. \\ & \quad \left. + (\lambda_1 P_+(t, E) - Q_+(t, E))e^{\alpha t} \right) \\ & \quad + m t^{m-1} (P_-(t, E)e^{-\alpha t} + P_0(t, E) + P_+(t, E)e^{\alpha t}) \\ & = \sum_{j=1}^r \sum_{\ell=0}^{k_j-1} c_{j\ell} t^\ell e^{(j-1)\alpha t}. \end{aligned} \quad (6.66)$$

At the right hand side we have terms with factors  $e^{\alpha t}$ ,  $e^{2\alpha t}$ ,  $\dots$ , but not with  $e^{-\alpha t}$ . So necessarily

- (i)  $\lambda_1 P_-(t, E) - Q_-(t, E) = 0$ ,
- (ii) if  $k_1 > 1$ , then  $P_-(t, E) = 0$ .

By considering the analogue of (6.66) for  $p = r$ , we similarly find

- (iii)  $\lambda_r P_+(t, E) - Q_+(t, E) = 0$ ,
- (iv) if  $k_r > 1$ , then  $P_+(t, E) = 0$ .

Next we consider the analogue of (6.66) for  $p = 2$  and concentrate on the terms with factor  $e^{-\alpha t}$ . At the right hand side  $e^{-\alpha t}$  is multiplied by a polynomial of degree  $k_1 - 1$  while at the left hand side  $m$  might go up to  $k_2 - 1$  and  $\lambda_2 P_-(t, E) - Q_-(t, E) \neq 0$  unless both  $P_-$  and  $Q_-$  are identically equal to zero. So if  $P_-$  and  $Q_-$  are nontrivial,  $k_1 = 1$  implies  $k_2 = 1$ . Likewise we may go down from  $r$  to  $r - 1$  and concentrate on terms with factor  $e^{\alpha t}$ . We find that if  $P_+$  and  $Q_+$  are nontrivial,  $k_r = 1$  implies  $k_{r-1} = 1$ . Continuing in this way we conclude that, unless  $P_-$ ,  $Q_-$ ,  $P_+$  and  $Q_+$  are all identically equal to zero, necessarily all  $k_i$ 's are equal to 1. It follows that  $P_-$ ,  $Q_-$ ,  $P_0$ ,  $Q_0$ ,  $P_+$  and  $Q_+$  all have degree zero, that is, they are constant functions of  $t$ .

Finally (6.65) follows from (6.26) and (6.31).  $\square$

So far it has been tacitly understood that  $g$ ,  $\mu$  and  $w$  are real valued functions and that the matrix  $H(E_0)$  has real entries. But the fact that  $H(E_0)$  is a real matrix does not preclude complex eigenvalues, it only implies that non-real eigenvalues occur in complex conjugate pairs. If we allow  $\alpha$  to be complex, the relation (6.42) tells us that the eigenvalues lie on a straight line in the complex plane. This is compatible with the requirement of eigenvalues occurring in complex conjugate pairs if and only if either  $\alpha$  is real or  $\alpha = i\beta$  for some real  $\beta$ . We conclude that either  $\alpha$  is real and all eigenvalues are real or  $\alpha = i\beta$  and the eigenvalues lie on a line parallel to the imaginary axis.

**Corollary 6.12** *Let  $r \geq 2$ . If*

$$\bigcap_{p_1 \neq p_2} (U_{p_1} \cup U_{p_2})$$

*contains an  $\alpha \in \mathbb{R} \setminus \{0\}$ , then  $g$ ,  $\mu$ ,  $w$  and  $H$  have the form specified in  $F_3$ .*

**Proof** Combining (6.30) with Lemma 6.11 we find that

$$\begin{aligned} g(x, E) &= g(x, E_0) \left( u_1(E) e^{-\alpha\tau(x)} + u_2(E) + u_3(E) e^{\alpha\tau(x)} \right) \\ &= \frac{1}{\alpha} g(x, E_0) e^{-\alpha\tau(x)} \left( \alpha u_1(E) + \alpha u_2(E) e^{\alpha\tau(x)} + \alpha u_3(E) \left( e^{\alpha\tau(x)} \right)^2 \right) \end{aligned}$$

and thus  $g$  has the form specified in  $F_3$  with

$$\begin{aligned} v_0(x) &= \frac{1}{\alpha} g(x, E_0) e^{-\alpha\tau(x)}, \\ v_i(E) &= \alpha u_i(E), \quad i = 1, 2, 3, \\ \zeta(x) &= e^{\alpha\tau(x)}. \end{aligned} \tag{6.67}$$

Note that by (6.27)

$$\zeta'(x) = \alpha \frac{1}{g(x, E_0)} e^{\alpha\tau(x)} = \frac{1}{v_0(x)},$$

so

$$\zeta(x) = \int_{x_b}^x \frac{dy}{v_0(y)}$$

as it should.

From (6.25) and (6.64) it follows that

$$\begin{aligned} \mu(x, E) &= \frac{\mu(x, E_0)}{g(x, E_0)} g(x, E) + \lambda_1 \left( u_1(E) e^{-\alpha\tau(x)} + u_3(E) e^{\alpha\tau(x)} \right) \\ &\quad + u_4(E) + (k-1) \alpha u_3(E) e^{\alpha\tau(x)} \\ &= \gamma_0(x) g(x, E) + \mu_0(E) + (k-1) v_3(E) \zeta(x), \end{aligned}$$

where, in addition to (6.67), we have defined

$$\begin{aligned} \gamma_0(x) &= \frac{\mu(x, E_0) + \lambda_1}{g(x, E_0)}, \\ \mu_0(E) &= u_4(E) - \lambda_1 u_2(E). \end{aligned}$$

Now recall (6.32) and (6.42). Since all eigenvalues of  $H(E_0)$  have multiplicity one, we can choose for  $H(E_0)$  the diagonal matrix with the eigenvalues on the diagonal. With the choice  $\Phi_j(0) = 1$ ,  $j = 1, 2, \dots, k$ , we have

$$\Phi_j(t) = e^{\lambda_1 t} e^{(j-1)\alpha t}.$$

A combination of (6.23) and (6.28) next yields

$$w(x) = e^{\int_{x_b}^x \frac{\mu(y, E_0)}{g(y, E_0)} dy} \Phi(\tau(x)) \quad (6.68)$$

and the definitions of  $\zeta$ ,  $\gamma_0$  and  $\Phi_j$  guarantee that (6.68) is exactly of the form specified in  $F_3$ .

Plugging in the information we now have about  $g$ ,  $\mu$  and  $w$  into (4.33), we find that  $H(E)$  has the form specified in  $F_3$ . Note that (6.65) of Lemma 6.11 implies that  $v_1(E_0) = \alpha u_1(E_0) = 0$ ,  $v_2(E_0) = \alpha u_2(E_0) = \alpha$  and  $v_3(E_0) = \alpha u_3(E_0) = 0$  which shows that  $H(E_0)$  is indeed the diagonal matrix specified above.  $\square$

When

$$\bigcap_{p_1 \neq p_2} (U_{p_1} \cup U_{p_2})$$

contains a purely imaginary number  $\alpha = i\beta$  ( $\beta$  real and  $\beta \neq 0$ ), we have  $\lambda_k = \bar{\lambda}_1$  and hence  $\operatorname{Im} \lambda_1 = -\frac{k-1}{2}\beta$ . It follows that

$$\lambda_j = \operatorname{Re} \lambda_1 + \left(j - \frac{k+1}{2}\right) i\beta. \quad (6.69)$$

In (6.63) we now allow  $u_1(E)$  to be complex while requiring  $u_3(E) = \overline{u_1(E)}$ . Using (6.30) we find

$$\begin{aligned} g(x, E) &= g(x, E_0) \left( u_1(E) e^{-i\beta\tau(x)} + u_2(E) + \overline{u_1(E)} e^{i\beta\tau(x)} \right) \\ &= g(x, E_0) (c_1(E) \cos \beta\tau(x) + c_2(E) \sin \beta\tau(x) + u_2(E)), \end{aligned} \quad (6.70)$$

where

$$\begin{aligned} c_1(E) &:= 2 \operatorname{Re} u_1(E), \\ c_2(E) &:= 2 \operatorname{Im} u_1(E). \end{aligned} \quad (6.71)$$

The well-known trigonometric identities

$$\begin{aligned} \cos \theta &= \cos^2 \frac{1}{2}\theta - \sin^2 \frac{1}{2}\theta \\ \sin \theta &= 2 \sin \frac{1}{2}\theta \cos \frac{1}{2}\theta \\ 1 &= \cos^2 \frac{1}{2}\theta + \sin^2 \frac{1}{2}\theta \\ \frac{d}{d\theta} \tan \theta &= \frac{1}{\cos^2 \theta} \end{aligned}$$

allow us to rewrite (6.70) as

$$g(x, E) = g(x, E_0) \frac{\cos^2 \frac{1}{2} \beta \tau(x)}{\frac{1}{2} \beta} \left( \frac{1}{2} \beta (u_2(E) + c_1(E)) \right. \\ \left. + \beta c_2(E) \tan \frac{1}{2} \beta \tau(x) + \frac{1}{2} \beta (u_2(E) - c_1(E)) \left( \tan \frac{1}{2} \beta \tau(x) \right)^2 \right), \quad (6.72)$$

which is of the form specified in  $F_3$  with

$$\zeta(x) = \tan \frac{1}{2} \beta \tau(x) \quad (6.73)$$

because

$$\zeta'(x) = \frac{\frac{1}{2} \beta \tau'(x)}{\cos^2 \frac{1}{2} \beta \tau(x)} = \frac{\frac{1}{2} \beta}{g(x, E_0) \cos^2 \frac{1}{2} \beta \tau(x)}.$$

Using  $\alpha = i\beta$  and (6.69) we rewrite (6.64) as

$$\tilde{\mu}(\tau^{-1}(t), E) = d_1(E) \cos \beta t + d_2(E) \sin \beta t + u_5(E) + \operatorname{Re} \lambda_1 \tilde{g}(t, E)$$

with

$$u_5(E) := u_4(E) - \operatorname{Re} \lambda_1 u_2(E), \\ d_1(E) := \frac{k-1}{2} \beta c_2(E), \\ d_2(E) := -\frac{k-1}{2} \beta c_1(E). \quad (6.74)$$

Equivalently we can write

$$\tilde{\mu}(\tau^{-1}(t), E) \\ = \cos^2 \frac{1}{2} \beta t \left( u_5(E) + d_1(E) + 2d_2(E) \tan \frac{1}{2} \beta t \right. \\ \left. + (u_5(E) - d_1(E)) \left( \tan \frac{1}{2} \beta t \right)^2 \right) \\ + \operatorname{Re} \lambda_1 \tilde{g}(t, E),$$

which, in combination with (6.25) and (6.30) yields

$$\mu(x, E) = \frac{\mu(x, E_0) + \operatorname{Re} \lambda_1}{g(x, E_0)} g(x, E)$$

$$\begin{aligned}
& + \frac{1}{1 + \zeta(x)^2} (u_5(E) + d_1(E) + 2d_2(E)\zeta(x) \\
& + (u_5(E) - d_1(E))\zeta(x)^2), \tag{6.75}
\end{aligned}$$

where  $\zeta$  is defined in (6.73). Thus our task is find  $\gamma_0(x)$  and  $\mu_0(E)$  such that

$$\mu(x, E) = \gamma_0(x)g(x, E) + \mu_0(E) + (k-1)\frac{1}{2}\beta(u_2(E) - c_1(E))\zeta(x). \tag{6.76}$$

The right hand sides of (6.75) and (6.76) are equal if and only

$$\begin{aligned}
& \frac{1}{1 + \zeta(x)^2} \left( u_5(E) + d_1(E) + 2d_2(E)\zeta(x) + (u_5(E) - d_1(E))\zeta(x)^2 \right) \\
& = \frac{\tilde{\gamma}_0(x)}{1 + \zeta(x)^2} \left( \frac{1}{2}\beta(u_2(E) + c_1(E)) + \beta c_2(E)\zeta(x) \right. \\
& \quad \left. + \frac{1}{2}\beta(u_2(E) - c_1(E))\zeta(x)^2 \right) \\
& \quad + \mu_0(E) + (k-1)\frac{1}{2}\beta(u_2(E) - c_1(E))\zeta(x), \tag{6.77}
\end{aligned}$$

where we have replaced the unknown function  $\gamma_0$  by a new unknown function  $\tilde{\gamma}_0$  via the relation

$$\gamma_0(x) = \frac{\mu(x, E_0) + \operatorname{Re} \lambda_1}{g(x, E_0)} + \frac{\frac{1}{2}\beta\tilde{\gamma}_0(x)}{g(x, E_0)} \tag{6.78}$$

and where we used (6.72), (6.73) and the identity  $(1 + \zeta(x)^2) \cos^2 \frac{1}{2}\beta\tau(x) = 1$ .

We rewrite the left hand side of (6.77) as

$$u_5(E) + d_1(E) + \frac{2\zeta(x)}{1 + \zeta(x)^2} (d_2(E) - d_1(E)\zeta(x))$$

and choose  $\mu_0(E) = u_5(E) + d_1(E)$ . This leaves us with

$$\begin{aligned}
& \frac{2\zeta(x)}{1 + \zeta(x)^2} (d_2(E) - d_1(E)\zeta(x)) \\
& = \frac{\tilde{\gamma}_0(x)}{1 + \zeta(x)^2} \left( \frac{1}{2}\beta(u_2(E) + c_1(E)) + \beta c_2(E)\zeta(x) + \frac{1}{2}\beta(u_2(E) - c_1(E))\zeta(x)^2 \right) \\
& \quad + (k-1)\frac{1}{2}\beta(u_2(E) - c_1(E))\zeta(x).
\end{aligned}$$

Hence we should have

$$\frac{1}{\zeta(x)}\tilde{\gamma}_0(x) = \frac{2d_2(E) - 2d_1(E)\zeta(x) - (k-1)\frac{1}{2}\beta(u_2(E) - c_1(E))(1 + \zeta(x)^2)}{\frac{1}{2}\beta(u_2(E) + c_1(E)) + \beta c_2(E)\zeta(x) + \frac{1}{2}\beta(u_2(E) - c_1(E))\zeta(x)^2}$$

$$\begin{aligned}
&= \frac{2d_2(E) - (k-1)\frac{1}{2}\beta(u_2(E) - c_1(E)) - 2d_1(E)\zeta(x)}{\frac{1}{2}\beta(u_2(E) + c_1(E)) + \beta c_2(E)\zeta(x) + \frac{1}{2}\beta(u_2(E) - c_1(E))\zeta(x)^2} \\
&\quad - \frac{(k-1)\frac{1}{2}\beta(u_2(E) - c_1(E))\zeta(x)^2}{\frac{1}{2}\beta(u_2(E) + c_1(E)) + \beta c_2(E)\zeta(x) + \frac{1}{2}\beta(u_2(E) - c_1(E))\zeta(x)^2} \\
&= -(k-1),
\end{aligned}$$

where the last equality is a consequence of (6.74). We conclude that  $\mu$  is of the form specified in  $F_3$  with

$$\begin{aligned}
\mu_0(E) &= u_5(E) + d_1(E) \\
\gamma_0(x) &= \frac{\mu(x, E_0) + \operatorname{Re} \lambda_1 - \frac{1}{2}\beta(k-1)\zeta(x)}{g(x, E_0)}. \tag{6.79}
\end{aligned}$$

The next task is to specify a *real* matrix  $H(E_0)$  with eigenvalues  $\lambda_j$  given by (6.69). From (6.72) we deduce that necessarily  $c_1(E_0) = c_2(E_0) = 0$  and  $u_2(E_0) = 1$ . This translates into  $v_2(E_0) = 0$ ,  $v_1(E_0) = v_3(E_0) = \frac{1}{2}\beta$  for the parameters  $v_i$ ,  $i = 1, 2, 3$  in  $F_3$ . So if we put  $H(E_0) = H_0(E_0) - \operatorname{Re} \lambda_1 I$ , then the formulation of  $F_3$  suggests to consider

$$H_0(E_0) = \frac{1}{2}\beta M, \tag{6.80}$$

where  $M$  is the  $k \times k$ -matrix with entries

$$M_{\ell j} = \begin{cases} -(k-\ell) & \text{if } j = \ell + 1, \\ j & \text{if } \ell = j + 1, \\ 0 & \text{otherwise.} \end{cases} \tag{6.81}$$

In the present context the key point is to find a solution to the linear ODE system corresponding to  $M$ . The result is given in the following lemma.

**Lemma 6.13** Define the  $k$ -vector  $\tilde{\Phi}(t)$  by

$$\tilde{\Phi}_j(t) = \sin^{j-1} t \cos^{k-j} t, \quad j = 1, 2, \dots, k. \tag{6.82}$$

Then

$$\frac{d}{dt} \tilde{\Phi}(t) = M \tilde{\Phi}(t). \tag{6.83}$$

**Proof** The  $j$ th component of both sides equals

$$(j-1)\tilde{\Phi}_{j-1}(t) - (k-j)\tilde{\Phi}_{j+1}(t)$$

(with the convention that zero times the undefined  $\tilde{\Phi}_0$  or  $\tilde{\Phi}_{k+1}$  equals zero). □

In order to determine the eigenvalues of  $M$  we shall relate  $M$  to the Kac-Sylvester matrix  $K$  defined by

$$K_{\ell j} = \begin{cases} \ell & \text{if } j = \ell + 1, \\ k - j & \text{if } \ell = j + 1, \\ 0 & \text{otherwise.} \end{cases} \quad (6.84)$$

It is known (Taussky and Todd 1991) that the eigenvalues of  $K$  are given by

$$\lambda_j^K = 2j - k - 1, \quad j = 1, 2, \dots, k. \quad (6.85)$$

In fact much more is known and we recommend the reader to have a look at the paper by Taussky and Todd (1991).

**Lemma 6.14** *Let  $S$  be the  $k \times k$  diagonal matrix with complex diagonal entries*

$$S_{jj} = i^{j-1}, \quad j = 1, 2, \dots, k. \quad (6.86)$$

*Then*

$$SMS^{-1} = iK^T. \quad (6.87)$$

**Proof**  $S^{-1}$  is the diagonal matrix with entries

$$(S^{-1})_{jj} = (-i)^{j-1}$$

and hence

$$(SMS^{-1})_{\ell j} = \sum_{m=1}^k S_{\ell m} \sum_{n=1}^k M_{mn} (S^{-1})_{nj} = i^{\ell-1} M_{\ell j} (-i)^{j-1}.$$

If  $j = \ell + 1$ , then  $i^{\ell-1}(-i)^{j-1} = i^{-1} = -i$  and accordingly

$$(SMS^{-1})_{\ell j} = i(k - \ell) = iK_{j\ell}.$$

If  $\ell = j + 1$ , then  $i^{\ell-1}(-i)^{j-1} = (-i)^{-1} = i$  and in this case

$$(SMS^{-1})_{\ell j} = ij = iK_{j\ell}.$$

□

**Corollary 6.15** *The eigenvalues of  $M$  are given by*

$$\lambda_j^M = (2j - k - 1)i, \quad j = 1, 2, \dots, k. \quad (6.88)$$



Motivated by (6.69), (6.80) and Lemma 6.13 we now choose

$$H(E_0) = H_0(E_0) + \operatorname{Re} \lambda_1 I = \frac{1}{2} \beta M + \operatorname{Re} \lambda_1 I \quad (6.89)$$

and

$$\Phi(t) = \tilde{\Phi} \left( \frac{1}{2} \beta t \right) e^{\operatorname{Re} \lambda_1 t}. \quad (6.90)$$

Combining (6.90) with (6.23) and (6.28) we obtain

$$w(x) = e^{\int_{x_b}^x \frac{\mu(y, E_0)}{g(y, E_0)} dy} \tilde{\Phi} \left( \frac{1}{2} \beta \tau(x) \right) e^{\operatorname{Re} \lambda_1 \tau(x)}. \quad (6.91)$$

Since  $v_2(E_0) = 0$  and  $v_1(E_0) = v_3(E_0) = \frac{1}{2} \beta$  we have

$$g(x, E_0) = v_0(x) \frac{1}{2} \beta \left( 1 + \zeta(x)^2 \right) = \frac{1}{2} \beta \frac{1 + \zeta(x)^2}{\zeta'(x)}.$$

To see that

$$\mu(x, E_0) = \gamma_0(x) g(x, E_0) - \operatorname{Re} \lambda_1 + \frac{1}{2} \beta (k-1) \zeta(x)$$

we first deduce from (6.79) that

$$\mu_0(E_0) = u_5(E_0) + d_1(E_0).$$

next from (6.74) that  $d_1(E_0) = 0$  and

$$u_5(E_0) = u_4(E_0) - \operatorname{Re} \lambda_1 u_2(E_0)$$

and from (6.64) and (6.26) that  $u_4(E_0) = 0$  and, as above, from (6.72) that  $u_2(E_0) = 1$ .

So

$$\begin{aligned} \int_{x_b}^x \frac{\mu(y, E_0)}{g(y, E_0)} dy &= \int_{x_b}^x \gamma_0(y) dy + \int_{x_b}^x \frac{-\operatorname{Re} \lambda_1 + \frac{1}{2} \beta (k-1) \zeta(y)}{\frac{1}{2} \beta (1 + \zeta(y))^2} \zeta'(y) dy \\ &= \int_{x_b}^x \gamma_0(y) dy - \operatorname{Re} \lambda_1 \tau(x) + \frac{k-1}{2} \log \left( 1 + \zeta(x)^2 \right), \end{aligned}$$

where we have used the fact that, on account of (6.73),

$$\frac{\zeta'(y)}{1 + \zeta(y)^2} = \frac{1}{2} \beta \tau'(y)$$

and that

$$\frac{\zeta(y)\zeta'(y)}{1+\zeta(y)^2} = \frac{d}{dy} \frac{1}{2} \log(1+\zeta(y)^2).$$

Hence (6.91) can be written in the form

$$w(x) = e^{\int_{x_b}^x \gamma_0(y) dy} \left(1 + \zeta(x)^2\right)^{\frac{k-1}{2}} \tilde{\Phi}\left(\frac{1}{2}\beta\tau(x)\right)$$

and, using (6.73) once more, as

$$w(x) = e^{\int_{x_b}^x \gamma_0(y) dy} \frac{1}{\cos^{k-1} \frac{1}{2}\beta\tau(x)} \tilde{\Phi}\left(\frac{1}{2}\beta\tau(x)\right).$$

Recalling (6.82) we conclude that

$$w_j(x) = e^{\int_{x_b}^x \gamma_0(y) dy} \left(\tan \frac{1}{2}\beta\tau(x)\right)^{j-1} \quad (6.92)$$

exactly as specified in  $F_3$  when  $\zeta$  is given by (6.73).

Finally we observe that once  $g(x, E)$ ,  $\mu(x, E)$  and  $w(x)$  have been determined, we choose  $H(E)$  such that (4.33) is indeed an identity. This leads to the matrix  $H(E)$  as specified in  $F_3$ .

We have thus reached the end of the proof of Theorem 5.3.

## 7 Applications to age- and stage-structured population models

In this section we illustrate our general results by giving a characterisation of ODE-reducible age-structured models and by giving conditions for when the ODE-reduction of a structured model could be interpreted biologically in terms of physiological stages.

Age is a peculiar i-state. It is characterised by two properties: age advances equally with time, so the growth rate  $g$  is identically equal to 1, and all individuals are born with the same i-state  $x_b = 0$ . The latter property means that the reproduction rate has the form

$$\beta(x, E, \cdot) = \tilde{\beta}(x, E)\delta_0, \quad (7.1)$$

that is, its range in the space of measures is the one-dimensional subspace spanned by  $\delta_0$ .

Incidentally, please note that these two properties guarantee that the p-state has a density  $n(t, \cdot)$  at time  $t > s$  if it has a density at time  $s$  and that the evolution in time of this density is governed by the PDE

$$\frac{\partial}{\partial t} n(t, x) + \frac{\partial}{\partial x} n(t, x) = -\mu(x, E(t))n(t, x) \quad (7.2)$$

with boundary condition

$$n(t, 0) = \int_0^\infty \tilde{\beta}(x, E(t))n(t, x)dx. \quad (7.3)$$

So here we are, in the spirit of Gurtin and MacCamy (1974, 1979a, b), Murphy (1983) and many others, classifying when the problem (7.2) and (7.3) reduces to a system of ODEs, but with two additional features:

- we consider explicitly output, that is, we specify what properties of the solutions matter;
- we want to derive conditions that are not only sufficient, but in fact necessary if reproduction is part of the output.

As explained in Sect. 4, the structured population model is ODE-reducible if the transport-degradation model is ODE-reducible and the reproduction process is such that there is a  $k \times k$  matrix  $M(E)$  such that (4.28) holds. The one-dimensional range property leads to simplifications, which we will now describe.

With (7.1) the restriction (4.28) becomes

$$\tilde{\beta}(x, E)w(0) = M(E)w(x), \quad (7.4)$$

or written out for each row  $i = 1, 2, \dots, k$ ,

$$\tilde{\beta}(x, E)w_i(0) = M_{i1}(E)w_1(x) + M_{i2}(E)w_2(x) + \dots + M_{ik}(E)w_k(x). \quad (7.5)$$

It follows that

$$\tilde{\beta}(x, E) = \sum_{j=1}^k \tilde{\beta}_j(E)w_j(x) \quad (7.6)$$

with

$$\tilde{\beta}_j(E) = \frac{M_{ij}(E)}{w_i(0)} \quad (7.7)$$

independent of  $i$ , that is, the elements  $M_{ij}(E)$  of the matrix  $M(E)$  are

$$M_{ij}(E) = w_i(0)\tilde{\beta}_j(E) \quad (7.8)$$

and thus  $M(E)$ , too, has one-dimensional range (namely the subspace spanned by  $(w_1(0), w_2(0), \dots, w_k(0))^T$ ).

Next we show that for age structured models the family  $F_3$  is included in family  $F_2$ .

Because  $g = 1$  in age-structured models, the entries for  $g$  in the families  $F_2$  and  $F_3$  imply that the functions  $v_1$ ,  $v_2$  and  $v_3$  are constants. In the case of family  $F_2$  it follows

that also  $v_0$  is a constant, which we without loss of generality take equal to 1 and so

$$\zeta(x) = \int_0^x \frac{dy}{v_0(y)} = x. \quad (7.9)$$

On the other hand, in the case of family  $F_3$ , the entry for  $g$  and the fact that  $\zeta'(x) = 1/v_0(x)$  and  $g = 1$  imply that  $\zeta(x)$  is the unique solution of the initial value problem

$$\zeta'(x) = v_1 + v_2\zeta(x) + v_3\zeta(x)^2, \quad \zeta(0) = 0. \quad (7.10)$$

We now deduce from the entries for  $\mu$  that in all three families the death rate  $\mu(x, E)$  is a sum of two functions, one depending only on  $x$  and the other depending only on  $E$ :

$$\mu(x, E) = \gamma_0(x) + \mu_0(E), \quad (7.11)$$

in the case of family  $F_2$  and

$$\mu(x, E) = (\tilde{\gamma}_0(x) + (k-1)v_3\zeta(x)) + \mu_0(E), \quad (7.12)$$

in the case of family  $F_3$ . In (7.12) we have added a tilde to the  $\gamma_0$  parameter in family  $F_3$ . The reason for this will become clear soon.

The next lemma gives the relationship between the families  $F_2$  and  $F_3$ .

**Lemma 7.1** *The  $F_2$  member with  $g = 1$ ,  $\zeta(x) = x$ ,  $\Lambda = H_0$  and  $\mu(x, E)$  given by (7.11) coincides with the  $F_3$  member with  $g = 1$ ,  $\zeta$  the solution of (7.10) and  $\mu(x, E)$  given by (7.12) with*

$$\tilde{\gamma}_0(x) = \gamma_0(x) - (k-1)v_3\zeta(x). \quad (7.13)$$

**Proof** Take  $w$  from the family  $F_3$ :

$$w_j(x) = e^{\int_0^x \tilde{\gamma}_0(y) dy} \zeta(x)^{j-1}, \quad j = 1, 2, \dots, k.$$

Then

$$\begin{aligned} w'_j(x) &= \tilde{\gamma}_0(x)w_j(x) + (j-1)w_{j-1}(x)\zeta'(x) \\ &= \gamma_0(x)w_j(x) + (j-1)w_{j-1}(x)(v_1 + v_2\zeta(x)) \\ &\quad - (k-1)v_3\zeta(x)w_j(x) + (j-1)v_3w_{j-1}(x)\zeta(x)^2. \end{aligned} \quad (7.14)$$

Now note that

$$\zeta(x)w_j(x) = \zeta(x)^2w_{j-1}(x)$$

and that this quantity equals  $w_{j+1}(x)$  if  $j \leq k-1$ . For  $j = k$  the last two terms of (7.14) cancel. This shows that  $w$  satisfies the equation

$$w'(x) = (\gamma_0(x)I + H_0)w(x),$$

that is,  $w$  is from family  $F_2$  with  $\Lambda = H_0$ .  $\square$

By Lemma 7.1 it suffices to concentrate on family  $F_2$  of the catalogue (the scalar representation of family  $F_1$  is a very special case of this). We have the following result:

**Theorem 7.2** *Under the restriction that the birth rate is of the form (7.6), an age-structured population model is ODE-reducible if and only if the death rate can be written as (7.11) and the  $w$  occurring in (7.6) has the form*

$$w(x) = e^{\int_0^x \gamma_0(y)dy} e^{x\Lambda} w(0) \quad (7.15)$$

for some  $k \times k$  matrix  $\Lambda$  and  $w(0)$  satisfying (H2) and (H3).

When this is the case,

$$N(t) = \int_0^\infty w(x)m(t, dx) = \int_0^\infty w(x)n(t, x)dx \quad (7.16)$$

satisfies the ODE

$$\frac{d}{dt}N(t) = (M(E(t)) + \Lambda - \mu_0(E(t)))N(t), \quad (7.17)$$

with  $M(E)$  defined by (7.8).

Once  $N$  has been solved from (7.17), the population density is recovered from the explicit formula

$$n(t, x) = e^{-\int_0^x \gamma_0(y)dy} e^{-\int_{t-x}^t \mu_0(E(\tau))d\tau} \sum_{j=1}^k \tilde{\beta}_j(E(t-x))N_j(t-x), \quad x < t. \quad (7.18)$$

Age-structured models can also be written directly as scalar renewal equations (with input dependent kernel) for the population birth rate. In an earlier paper (Diekmann et al. 2018) we derived necessary and sufficient conditions for ODE-reducibility of such renewal equations. When applied to age-structured models the results of the two papers are of course the same.

Let us point out that the ODE (7.17) resulting from the reduction does not necessarily have any biological interpretation. But if  $\Lambda$  is a transition matrix, then the indices of  $N$  could be interpreted as discrete states. In particular, if  $\Lambda$  has negative entries on the diagonal and corresponding positive entries on the subdiagonal, then the states form a linear chain and the states can be interpreted as stages (Nisbet and Gurney 1983;

Nisbet et al. 1985, 1986; Thieme 2003) if, in addition,  $w_j(0) = 0$  for  $j \geq 2$ , which means that all newborns have stage  $j = 1$ . This explains why MacDonald (1978, 1989) called ODE-reduction the *linear chain trick*. The classical example of such linear chain trickery is obtained by choosing  $\gamma_0 = 0$ , and  $w_j$  the Erlang distribution (a special case of the gamma distribution)

$$w_j(x) = \frac{\alpha^j}{(j-1)!} x^{j-1} e^{-\alpha x}, \quad (7.19)$$

and  $A$  a matrix with diagonal elements  $-\alpha$  and subdiagonal elements  $+\alpha$ . The ODE (7.17) then becomes

$$\frac{d}{dt} N_1(t) = \alpha \sum_{j=1}^k \tilde{\beta}_j(E) N_j - (\alpha + \mu_0(E)) N_1, \quad (7.20)$$

$$\frac{d}{dt} N_i(t) = \alpha N_{i-1} - (\alpha + \mu_0(E)) N_i, \quad i = 2, 3, \dots, k, \quad (7.21)$$

clearly showing the stage structure (Gyllenberg 2007).

## 8 Discussion

Our focus should be on the biological problem: when is a physiologically structured model ODE representable? That is the question that matters, both practically and foundationally, for community modelling. In the wake of an answer we can re-examine the community models that are directly formulated in terms of ODE in order to determine what i-level physiological processes are, or are not, incorporated.

If we are prepared to specify beforehand a set of  $k_0$  linearly independent output functionals  $P_i^0$  that should be computable from the ODE variables, we can perform a test. The test was explained in Sects. 2.1.4 and 4.1, but we repeat it here while employing, exactly as in (2.18), the short hand of an abstract ODE

$$\frac{d}{dt} n(t) = A(E(t))n(t) \quad (8.1)$$

rather than the constructively defined evolutionary system  $U_E^c(t, s)$  of Sect. 3. Note that now we do *not* use the decomposition  $A(E) = A_0(E) + B(E)$ .

TEST: Let  $w_i^0$  be the element of the dual space representing  $P_i^0$ . A first condition for ODE representability is that  $w_i^0$  belongs to the domain of  $A(E)^*$  for all  $E$  in  $\mathcal{E}$ . If  $A(E)^* w_i^0$  belongs to the subspace spanned by  $\{w_1^0, w_2^0, \dots, w_{k_0}^0\}$  for all  $i$  and all  $E$ , the family  $A(E)$  is ODE-reducible. If the subspace spanned by  $\{w_1^0, w_2^0, \dots, w_{k_0}^0\}$  and the  $A(E)^* w_i^0$  for all  $i$  and all  $E$  is infinite dimensional, the family  $A(E)$  is not ODE-reducible. In case this subspace has dimension  $k_1$ , with  $k_1$  finite but larger than  $k_0$ , choose a basis consisting of  $w_i^1$  with  $i = 1, \dots, k_1$  and  $w_i^1 = w_i^0$  for  $i = 1, 2, \dots, k_0$  and repeat this procedure with the superscript 0 replaced by the 1. The procedure can

be repeated. If the procedure terminates after a finite number of steps with a finite dimensional subspace, the family  $A(E)$  is ODE-reducible.

Whenever biological knowledge and interest suggest a meaningful collection of  $w_i^0$ , one should perform this test. Clearly this requires context specific information and does not lead to an explicit catalogue. In this paper we concentrated on a special subclass of physiologically structured models for which we could obtain detailed results. The description of the subclass is based on the decomposition  $A(E) = A_0(E) + B(E)$ , with  $A_0(E)$  capturing transport and degradation and  $B(E)$  capturing reproduction. Our main result provides a complete catalogue of ODE reducible families  $A_0(E)$ . The catalogue lists conditions on the growth rate  $g$  and the mortality rate  $\mu$  and provides explicit expressions for the functionals that define the components of the ODE. Once these functionals are known, one can test whether or not the subspace that they span is left invariant by  $B(E)^*$  for all  $E$  in  $\mathcal{E}$ . If the subspace is invariant, the family  $A(E)$  is ODE-reducible. Moreover, the asymptotic population dynamics is in that case fully determined by the ODE (recall (2.27) and Proposition 7.2).

So the catalogue is complete for a restricted class of structured population models. We now provide some examples of biologically relevant ODE-reducible models that do *not* belong to that particular restricted class, because the reproduction process does not satisfy condition (4.28). Our first example shows how the simple model of Example 1.1 can be treated.

**Example 8.1** Assume a one-dimensional i-state space and that  $\beta$  is of the form (4.32). Equation (4.27) then becomes

$$g(x, E)w'(x) - \mu(x, E)w(x) + \tilde{\beta}(x, E) \int_{\Omega} w(y)\beta_0(dy) = K(E)w(x). \quad (8.2)$$

For definiteness, we normalise  $\beta_0$  by

$$\int_{\Omega} \beta_0(dy) = 1. \quad (8.3)$$

If we are seeking a scalar representation ( $w$  is a scalar valued function), then the condition (8.2) transforms into

$$\frac{g(x, E)w'(x) + \tilde{\beta}(x, E) \int_{\Omega} w(y)\beta_0(dy)}{w(x)} - \mu(x, E) = K(E). \quad (8.4)$$

If the output is the total population, then  $w(x) = 1$  and (8.4) becomes

$$\tilde{\beta}(x, E) - \mu(x, E) = K(E). \quad (8.5)$$

If the output is total biomass, then  $w(x) = x$  and (8.4) becomes

$$\frac{g(x, E) + \tilde{\beta}(x, E) \int_{\Omega} y\beta_0(dy)}{x} - \mu(x, E) = K(E). \quad (8.6)$$

Note that the scalar factor  $\int_{\Omega} y \beta_0(dy)$  is the average weight of a newborn. In the case of one fixed size  $x_b$  at birth,  $\beta_0 = \delta_{x_b}$  and (8.6) becomes the same as condition (1.1).

We now provide some other examples that are ODE-reducible, but do not satisfy the restrictions of the present paper.

**Example 8.2** Let  $\Omega$  be a subset of the positive cone  $\mathbb{R}_+^2$  and let

$$g(x, E) = \begin{pmatrix} a(E) + b(E)x_1 \\ c(E) \end{pmatrix}. \quad (8.7)$$

Then

$$(Dw)(x)g(x, E) = L(E)w(x) \quad (8.8)$$

for the choice

$$L(E) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ a(E) & b(E) & 0 & 0 & 0 & 0 \\ 0 & 2a(E) & 2b(E) & 0 & 0 & 0 \\ 0 & 0 & 0 & -\kappa c(E) & 0 & 0 \\ 0 & 0 & 0 & a(E) & b(E) - \kappa c(E) & 0 \\ 0 & 0 & 0 & 0 & 2a(E) & 2b(E) - \kappa c(E) \end{pmatrix} \quad (8.9)$$

and

$$w(x) = \begin{pmatrix} 1 \\ x_1 \\ x_1^2 \\ e^{-\kappa x_2} \\ x_1 e^{-\kappa x_2} \\ x_1^2 e^{-\kappa x_2} \end{pmatrix}. \quad (8.10)$$

So if

$$\mu(x, E) = \mu_0(E) \quad (8.11)$$

and

$$\beta(x, E, \omega) = f(E) (1 - e^{-\kappa x_2}) x_1^2 m_b(\omega) \quad (8.12)$$

with  $m_b$  a probability measure on  $\Omega$ , then

$$\begin{aligned} (Dw(x))g(x, E) - \mu(x, E)w(x) + \int_{\Omega} w(y)\beta(x, E, dy) \\ = L(E)w(x) - \mu_0(E)w(x) + f(E)(w_3(x) - w_6(x)) \int_{\Omega} w(y)m_b(dy) \end{aligned} \quad (8.13)$$



and hence (4.27) is satisfied with

$$K(E) = L(E) - \mu_0(E)I + f(E)M, \quad (8.14)$$

where  $M$  is the  $6 \times 6$  matrix with entries

$$M_{ij} := \int_{\Omega} w_i(y) m_b(dy) (\delta_{3j} - \delta_{6j}). \quad (8.15)$$

Here  $\delta$  is the Kronecker  $\delta$ , that is,  $\delta_{ij} = 1$  if  $i = j$  and zero otherwise. In words:  $M$  is the matrix with only two non-zero columns, the third and the sixth, and on the  $i$ th row the element equals  $\int_{\Omega} w_i(y) m_b(dy)$  in column 3 and  $-\int_{\Omega} w_i(y) m_b(dy)$  in column 6.

The biological interest of Example 8.2 is that if we let the measure  $m_b$  be concentrated on the axis  $x_2 = 0$ , we may interpret  $x_1$  as size and  $x_2$  as physiological age. The rate of giving birth is proportional to the square of the individual's size with the proportionality coefficient a function of physiological age that increases monotonically from zero at zero to a finite limit depending on the environmental condition.

The next two examples are not biologically motivated, but are intended to illustrate that, even after scaling of the  $i$ -state variable, the growth rate is not restricted to a linear dependence on the  $i$ -state components (recall that in our catalogue only  $F_3$  allows for a quadratic dependence on the  $i$ -state variable, but only so if the death rate  $\mu$  contains a  $k$ -specific compensating term).

**Example 8.3** Let  $\Omega \subset \mathbb{R}^2$  and let

$$g(x, E) = \begin{pmatrix} a(E) + b(E)x_1 \\ c(E)x_1^2 \end{pmatrix}. \quad (8.16)$$

Then (8.8) holds for the choice

$$L(E) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ a(E) & b(E) & 0 & 0 \\ 0 & 2a(E) & 2b(E) & 0 \\ 0 & 0 & c(E) & 0 \end{pmatrix} \quad (8.17)$$

and

$$w(x) = \begin{pmatrix} 1 \\ x_1 \\ x_1^2 \\ x_2 \end{pmatrix}. \quad (8.18)$$

**Example 8.4** Let  $\Omega \subset \mathbb{R}^3$  and let

$$g(x, E) = \begin{pmatrix} a_1(E) \\ a_2(E) \\ c_1(E)e^{\kappa_1 x_1} + c_2(E)e^{\kappa_2 x_2} + c_3(E)e^{\kappa_1 x_1 + \kappa_2 x_2} \end{pmatrix}. \quad (8.19)$$

Then (8.8) holds for the choice

$$L(E) = \begin{pmatrix} \kappa_1 a_1(E) & 0 & 0 & 0 \\ 0 & \kappa_2 a_2(E) & 0 & 0 \\ 0 & 0 & \kappa_1 a_1(E) + \kappa_2 a_2(E) & 0 \\ c_1(E) & c_2(E) & c_3(E) & 0 \end{pmatrix} \quad (8.20)$$

and

$$w(x) = \begin{pmatrix} e^{\kappa_1 x_1} \\ e^{\kappa_2 x_2} \\ e^{\kappa_1 x_1 + \kappa_2 x_2} \\ x_3 \end{pmatrix}. \quad (8.21)$$

**Example 8.5** Consider cells that split into two equal parts upon reaching a size  $x_{\max}$ . Assume that  $g(x, E) = f(E)x$  and that  $\mu(x, E) = \mu_0(E)$ . Then the total mass  $B$  satisfies the ODE

$$\frac{dB}{dt} = f(E)B - \mu_0(E)B.$$

If we start with a cohort (a Dirac mass in a point somewhere in the interval  $(\frac{1}{2}x_{\max}, x_{\max})$ ), the distribution remains a cohort for all time, but the cohort jumps from  $x_{\max}$  to  $\frac{1}{2}x_{\max}$  every now and then, with the number of individuals in the cohort doubling whenever such a jump occurs (Huyer 1997). Note that when we change the reproduction rule and allow the two daughters to have slightly different sizes, the number of cohorts increases.

In an earlier paper (Diekmann et al. 2018) we gave a characterisation of ODE-reducibility for renewal equations of the form

$$b(t) = \int_{-\infty}^t K_E(t, s)b(s)ds, \quad (8.22)$$

where the unknown  $b$  is an  $\mathbb{R}^p$  valued function and the input-dependent kernel  $K_E$  is a  $p \times p$  matrix valued function. The result is that the renewal Eq. (8.22) with input is ODE-reducible if and only

$$K_E(t, s) = U(E(t))^T \Phi_E(t, s) V(E(s)), \quad (8.23)$$

where  $\Phi_E(t, s)$  is the fundamental matrix solution for the linear non-autonomous system of ordinary differential equations

$$\frac{d}{d\tau} Z(\tau) = H(E(\tau))Z(\tau) \quad (8.24)$$

for some matrix valued functions  $H$ ,  $U$  and  $V$ . It should, however, be noted that in the earlier paper the requirement that the solution  $b(t)$  of the renewal Eq. (8.22) could

be recovered from the solution  $Z(t)$  of the finite dimensional system (8.24) was built into the notion of ODE-reducibility, whereas we in the current paper only require that the *outputs* can be recovered (recall Fig. 2). More recently we have taken the abstract variant of (8.22) as the starting point for a discussion of ODE-reducibility. In the paper (Diekmann et al. 2019) we provide new examples, recap the present paper and explain the usefulness of *asymptotic* ODE-reducibility.

**Acknowledgements** Open access funding provided by University of Helsinki including Helsinki University Central Hospital. Parts of the manuscript were written while the first two authors enjoyed the hospitality of the Banach Centre in Warsaw during the Simons semester on Mathematical Biology. We thank Piotr Gwiazda for making this happen. The manuscript was completed while the three of us participated in the research Programme “Mathematical Biology” at the Institute Mittag-Leffler in the fall semester 2018. We thank the professional and friendly staff for providing such excellent working conditions. J.A.J.M. benefitted from the support from the “Chair Modélisation Mathématique et Biodiversité of Veolia Environnement-Ecole Polytechnique-Museum National d’Histoire Naturelle-Fondation X”. Finally we thank two referees and Editor Mark Lewis for constructive and very helpful feedback.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References


- Atay FM, Roncoroni L (2017) Lumpability of linear evolution equations in Banach spaces. *Evol Equ Control Theory* 6(1):15–34
- Bacaër N (2008) *Histoires de Mathématiques et de Populations*. Cassini, Paris
- Bacaër N (2011) *A short history of mathematical population dynamics*. Springer, London
- Bôcher M (1900) The theory of linear dependence. *Ann Math Second Ser* 2:81–96
- Clément P, Diekmann O, Gyllenberg M, Heijmans HJAM, Thieme HR (1988) Perturbation theory for dual semigroups. II. Time-dependent perturbations in the sun-reflexive case. *Proc R Soc Edinb* 109A:145–172
- Cushing J (1989) A competition model for size-structured species. *SIAM J Appl Math* 49:838–858
- de Roos AM, Persson L (2013) *Population and community ecology of ontogenetic development*. Princeton University Press, Princeton
- de Roos AM, Metz JAJ, Evers E, Leipoldt A (1990) A size dependent predator-prey interaction: who pursues whom? *J Math Biol* 28:609–643
- de Roos AM, Metz JAJ, Persson L (2013) Ontogenetic symmetry and asymmetry in energetics. *J Math Biol* 66(4–5):889–914
- Diekmann O, Metz JAJ (2010) How to lift a model for individual behaviour to the population level? *Philos Trans R Soc B* 365:3523–3530
- Diekmann O, Gyllenberg M, Thieme HR (1995) Perturbing evolutionary systems by cumulative outputs and step responses. *Differ Integral Equ* 8:1205–1244
- Diekmann O, Gyllenberg M, Metz JAJ, Thieme HR (1998) On the formulation and analysis of general deterministic structured population models. I. Linear theory. *J Math Biol* 36:349–388
- Diekmann O, Gyllenberg M, Huang H, Kirkilionis M, Metz JAJ, Thieme HR (2001) On the formulation and analysis of general deterministic structured population models. II. Nonlinear theory. *J Math Biol* 43:157–189
- Diekmann O, Gyllenberg M, Metz JAJ (2003) Steady state analysis of structured population models. *Theor Popul Biol* 63:309–338
- Diekmann O, Gyllenberg M, Metz JAJ, Nakaoka S, de Roos AM (2010) *Daphnia revisited: local stability and bifurcation theory for physiologically structured population models explained by way of an example*. *J Math Biol* 61:277–318. Erratum: *J Math Biol* (2017) 75, 259–261

- Diekmann O, Gyllenberg M, Metz JAJ (2018) Finite dimensional state representation of linear and nonlinear delay systems. *J Dyn Differ Equ* 30:1439–1467
- Diekmann O, Gyllenberg M, Metz JAJ (2019) On models of physiologically structured populations and their reduction to ordinary differential equations. *J Math Biol*. <https://doi.org/10.1007/s00285-019-01431-7>
- Euler L (1760) Recherches générales sur la mortalité et la multiplication du genre humain. *Mém Acad R Sci Bell Lett Berlin* XVI:144–164
- Fargue D (1973) Réductibilité des systèmes héréditaires à des systèmes dynamiques (régis par des équations différentielles ou aux dérivées partielles). *CR Acad Sci Paris Sér B* 277:471–473
- Fargue D (1974) Réductibilité des systèmes héréditaires. *Int J Non-Linear Mech* 9:331–338
- Gripenberg G, Londen S-O, Staffans O (1990) Volterra integral and functional equations. Cambridge University Press, Cambridge
- Gurtin ME, MacCamy RC (1974) Non-linear age-dependent population dynamics. *Arch Ration Mech Anal* 54:281–300
- Gurtin ME, MacCamy RC (1979a) Some simple examples for non-linear age-dependent population dynamics. *Math Biosci* 43:199–211
- Gurtin ME, MacCamy RC (1979b) Population dynamics with age-dependence. In: Knops RJ (ed) *Nonlinear analysis and mechanics. Heriot-Watt symposium III*. Pitman, Boston, pp 1–35
- Gwiazda P, Marciniak-Czochra A (2010) Structured population equations in metric spaces. *J Hyperb Differ Equ* 7:733–773
- Gyllenberg M (2007) Mathematical aspects of physiologically structured populations: the contributions of J. A. J. Metz. *J Biol Dyn* 1:3–44
- Hirsch MW, Smale S (1974) Differential equations, dynamical systems, and linear algebra. *Pure and Applied Mathematics*, vol 60. Academic Press, New York
- Huyer W (1997) On periodic cohort solutions of a size-structured population model. *J Math Biol* 35:908–934
- Kingsland SE (1995) Modeling nature: episodes in the history of population ecology. The University of Chicago Press, Chicago
- Kooijman SALM, Metz JAJ (1984) On the dynamics of chemically stressed populations: the deduction of population consequences from effects on individuals. *Ecotoxicol Environ Saf* 8:254–274
- MacDonald M (1978) Time lags in biological models. *Lecture Notes in Biomathematics*, vol 27. Springer, Berlin
- MacDonald M (1989) Biological delay systems. Cambridge University Press, Cambridge
- Metz JAJ, Diekmann O (1986) The dynamics of physiologically structured populations. *Lecture Notes in Biomathematics*, vol 68. Springer, Berlin, Heidelberg, New York, London, Paris, Tokyo
- Metz JAJ, Diekmann O (1991) Exact finite dimensional representations of models for physiologically structured populations. I. The abstract foundations of linear chain trickery. In: Goldstein JA, Kappel F, Schappacher W (eds) *Differential equations with applications in biology, physics and engineering. Lecture Notes in Pure and Applied Mathematics*. Marcel Dekker, New York, pp 269–289
- Metz JAJ, Tran VC (2013) Daphnias: from the individual based model to the large population equation. *J Math Biol* 66:915–933
- Murphy LF (1983) A nonlinear growth mechanism in size structured population dynamics. *J Theor Biol* 104:493–506
- Nisbet RM, Gurney WSC (1983) The systematic formulation of population models for insects with dynamically varying instar duration. *Theor Popul Biol* 23:114–35
- Nisbet RM, Gurney WSC, Blythe SP, Metz JAJ (1985) Stage structure models of populations with distinct growth and development processes. *IMA J Math Appl Biol Med* 2:57–68
- Nisbet RM, Gurney WSC, Blythe SP (1986) The systematic formulation of models of stage-structured populations. In: Metz JAJ, Diekmann O (eds) *The dynamics of physiologically structured populations. Lecture Notes in Biomathematics* 68, Springer-Verlag, Berlin, Heidelberg, New York, London, Paris, Tokyo, pp 474–494
- Perthame B (2007) Transport equations in biology. Birkhäuser, Basel
- Rudin W (1973) Functional analysis. McGraw-Hill, New York
- Scudo FM, Ziegler JR (1998) The golden age of theoretical ecology: 1923–1940. Springer, Berlin
- Taussky O, Todd J (1991) Another look at a matrix of Mark Kac. *Linear Algebra Appl* 150:341–360
- Thieme HR (2003) Mathematics in population biology. Princeton University Press, Princeton
- Tran VC (2008) Large population limit and time behaviour of a stochastic particle model describing an age-structured population. *ESAIM Probab Stat* 12:345–386

Vogel T (1965) *Théorie des systèmes évolutifs*. Gauthier-Villars, Paris

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Affiliations

Odo Diekmann<sup>1</sup> · Mats Gyllenberg<sup>2</sup>  · Johan A. J. Metz<sup>3,4</sup>

✉ Mats Gyllenberg  
mats.gyllenberg@helsinki.fi

Odo Diekmann  
O.Diekmann@uu.nl

Johan A. J. Metz  
J.A.J.Metz@biology.leidenuniv.nl

<sup>1</sup> Department of Mathematics, University of Utrecht, P.O. Box 80010, 3508 TA Utrecht, The Netherlands

<sup>2</sup> Department of Mathematics and Statistics, University of Helsinki, P.O. Box 68, 00014 Helsinki, Finland

<sup>3</sup> Mathematical Institute and Institute of Biology, Leiden University, 2333 CA Leiden, The Netherlands

<sup>4</sup> Evolution and Ecology Program, International Institute of Applied Systems Analysis, 2361 Laxenburg, Austria